

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comphen Ltd.

## OM protein - protein search, using sw model

Run on: March 23, 2005, 18:59:31 ; Search time 169 Seconds  
(without alignments)  
469.147 Million cell updates/sec

Title: US-10-087-190-3  
Perfect score: 1047  
Sequence: 1 MSKKKGLSAEKRTFMELF.....FEENKIDRFGIPEDFDYID 205

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneeqp19808:\*  
2: geneeqp19908:\*  
3: geneeqp20008:\*  
4: geneeqp20018:\*  
5: geneeqp20028:\*  
6: geneeqp20038:\*  
7: geneeqp20038:\*  
8: geneeqp20048:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1047	100.0	205	ADD84537	Ad84537 121P1 P
2	1047	100.0	205	ADJ70015	AdJ70015 Human hea
3	1047	100.0	205	ADM83851	Adm83851 Human can
4	1047	100.0	205	ADM83810	Adm83810 Human can
5	1047	100.0	205	ADM83812	Adm83812 Human can
6	1047	100.0	205	ADM83793	Adm83793 Human can
7	1047	100.0	205	ADM83811	Adm83811 Human can
8	1047	100.0	205	ADM83804	Adm83804 Human can
9	1047	100.0	219	ABP75541	Abp75541 Human sec
10	1036.5	99.0	206	ADM83834	Adm83834 Human can
11	1036.5	99.0	206	ADM83835	Adm83835 Human can
12	1011	96.6	198	ADM83814	Adm83814 Human can
13	975	93.1	190	AAW40043	Aaw40043 Human can
14	975	93.1	190	ADD84547	Add84547 121P1 v
15	975	93.1	190	ADM83857	Adm83857 Human can
16	975	93.1	190	ADM83859	Adm83859 Human can
17	975	93.1	190	ADM83803	Adm83803 Human can
18	975	93.1	190	ADM83809	Adm83809 Human can
19	975	93.1	190	ADM83858	Adm83858 Human can
20	948	90.5	205	ADM83813	Adm83813 Mouse hyp
21	945.5	90.3	190	ADD84545	Add84545 121P1 v
22	945.5	90.3	190	ADM83808	Adm83808 Human can
23	945.5	90.3	190	ADM83801	Adm83801 Human can
24	945.5	90.3	190	ADM83853	Adm83853 Human can
25	945.5	90.3	190	ADM83852	Adm83852 Human can

26	614	58.6	122	8	ADM83843	Adm83843 Human can
27	606	57.9	122	4	AAW41829	Aaw41829 Human pol
28	591	56.4	122	7	ADD84543	Add84543 121P1 v
29	591	56.4	122	8	ADM83799	Adm83799 Human can
30	591	56.4	122	8	ADM83845	Adm83845 Human can
31	591	56.4	122	8	ADM83807	Adm83807 Human can
32	591	56.4	122	8	ADM83844	Adm83844 Human can
33	584	55.8	119	7	ADD84541	Add84541 121P1 v
34	584	55.8	119	8	ADM84541	Adm84541 Human can
35	584	55.8	119	8	ADM83806	Adm83806 Human can
36	584	55.8	119	8	ADM83837	Adm83837 Human can
37	465	44.4	126	7	ADD84539	Add84539 121P1 v
38	465	44.4	126	8	ADM83836	Adm83836 Human can
39	465	44.4	126	8	ADM83795	Adm83795 Human can
40	465	44.4	126	8	ADM83805	Adm83805 Human can
41	349	33.3	200	8	ADM83815	Adm83815 Yeast hyp
42	329.5	31.5	182	4	ABG12341	Abg12341 Novel hum
43	287	27.4	79	4	ACG74669	Acg74669 Human col
44	255.5	24.4	218	5	ADH32846	Adh32846 Yeast smc
45	135.5	12.9	1281	8	ADP26646	Adp26646 Mouse dyn

## ALIGNMENTS

RESULT 1  
ADD84537  
ID ADD84537 standard; protein; 205 AA.  
XX  
XX ADD84537;  
XX  
AC 29-JAN-2004 (first entry)  
DT XX  
DE 121P1 protein.  
XX  
XX 121P1; 121P1 modulation; human; chromosome 4q; cytostatic;  
KM gene therapy; vaccine; cancer; immune response; immunisation.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200295009-A2.  
PN  
XX  
PD 28-NOV-2002.  
XX  
XX 28-FEB-2002; 2002WO-US006242.  
PF  
XX  
XX 05-MAR-2001; 2001US-00799250.  
PR  
XX  
XX (AGEN-) AGENSYS INC.  
PA  
XX  
XX Challa-Bid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;  
PI Jakovovits A;  
XX  
XX WPI; 2003-156757/15.  
DR N-PSDB; ADD84536.  
XX  
PT Composition comprising a substance that modulates the status of 121P1,  
PT useful in diagnosing, preventing, prognosticating or treating patients  
PT with cancer that expresses 121P1, such as breast, colon, ovarian or  
PT lung cancer.  
XX  
PS Claim 19; Fig 2A; 285pp; English.  
XX  
XX The present invention describes a composition (I) comprising a substance  
XX that modulates the status of 121P1 (gene and encoded protein), or a  
XX molecule that is modulated by 121P1, where the status of a cell that  
XX expresses 121P1 is modulated. The human 121P1 gene maps to chromosome  
XX 4q. (I) has cytostatic activity, and can be used in gene therapy, and in  
XX vaccines. The composition (I) can be used for diagnosing, preventing,  
XX prognosticating or treating patients with cancer that expresses 121P1,  
XX such as breast, colon, ovarian or lung cancer. The 121P1 gene or its  
XX fragment can be used to elicit a humoral or cellular immune response.  
CC 121P1 antibodies can be used in active or passive immunisation. 121P1

CC polynucleotides are useful as probes and primers for the amplification or  
 CC detection of 121P1 genes, as coding sequences for directing the  
 CC expression of 121P1 polypeptides, or as tools for modulating the  
 CC inhibiting the expression of 121P1 genes. The present sequence is used  
 CC in the exemplification of the present invention.  
 CC Sequence 205 AA;

Query Match 100.0%; Score 1047; DB 7; Length 205;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-88;  
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSAEKKRTMMEIFSETDVFOLKDEKIAPEKIGITMSVKEVQSLVDDGMV 60  
 DB 1 MSKKKGLSAEKKRTMMEIFSETDVFOLKDEKIAPEKIGITMSVKEVQSLVDDGMV 60  
 QY 61 DCERIGTSNYMAFPSPKALHARKHLEVEQSLSGSGQKASLOKSTIEKAKIGRCETER 120  
 DB 61 DCERIGTSNYMAFPSPKALHARKHLEVEQSLSGSGQKASLOKSTIEKAKIGRCETER 120  
 QY 121 TRLAKELSLDQREQLKAEVEKYDCDPQVEEIRQANKVAKAANRWTDNIFAKSWA 180  
 DB 121 TRLAKELSLDQREQLKAEVEKYDCDPQVEEIRQANKVAKAANRWTDNIFAKSWA 180  
 QY 181 KRKFGFEENKIDRTFTGIPEDFDYID 205  
 DB 181 KRKFGFEENKIDRTFTGIPEDFDYID 205

RESULT 2  
 ADJ70015  
 ID ADJ70015 standard; protein; 205 AA.  
 AC ADJ70015;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX

DE Human heat mitochondrial protein as a therapeutic target SeqID1821.  
 XX  
 KW mitochondrial; human; screening assay; diabetes mellitus;  
 KW Huntington's disease; osteoarthritis;  
 KW Leber's hereditary optic neuropathy; LHON;  
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;  
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
 KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;  
 KW osteopathic; ophthalmological; cytostatic.  
 OS Homo sapiens.  
 XX  
 PN WO2003087768-A2.  
 XX  
 PD 23-OCT-2003.  
 XX  
 PF 04-APR-2003; 2003WO-US010870.  
 XX  
 PR 12-APR-2002; 2002US-0372843P.  
 PR 17-JUN-2002; 2002US-038987P.  
 PR 20-SEP-2002; 2002US-0412418P.  
 XX  
 PA (MITO-) MITOKOR  
 PA (BUCK-) BUCK INST AGE RES.  
 XX  
 PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;  
 PI Warnock DE;  
 XX  
 DR WPI; 2003-845369/78.  
 XX

Identifying a mitochondrial target for drug screening assays and for  
 PT treating diseases associated with altered mitochondrial function,  
 PT comprises detecting a modified polypeptide in a sample and correlating  
 PT with the disease.  
 XX  
 PS Claim 1; SEQ ID NO 1821; 180bp; English.

XX This invention relates to novel mitochondrial targets that can be used  
 CC for therapeutic intervention in treating a disease associated with  
 CC altered mitochondrial function. Specifically, it refers to a method for  
 CC identifying proteins of the human heart mitochondrial proteome that are  
 CC useful for drug screening assays, as well as therapeutic targets. The  
 CC present invention describes a method for identifying such proteins that  
 CC can be used in the treatment of various diseases associated with altered  
 CC mitochondrial function including diabetes mellitus, Huntington's disease,  
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
 CC encephalopathy, lactic acidosis and stroke (MELAS), myoclonic epilepsy  
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
 CC compositions have neuroprotective, nootropic, antidiabetic,  
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
 CC cyostatic activities. This polypeptide sequence is a human heart  
 CC mitochondrial protein of the invention.  
 CC Sequence 205 AA;

Query Match 100.0%; Score 1047; DB 7; Length 205;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-88;  
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSAEKKRTMMEIFSETDVFOLKDEKIAPEKIGITMSVKEVQSLVDDGMV 60  
 DB 1 MSKKKGLSAEKKRTMMEIFSETDVFOLKDEKIAPEKIGITMSVKEVQSLVDDGMV 60  
 QY 61 DCERIGTSNYMAFPSPKALHARKHLEVEQSLSGSGQKASLOKSTIEKAKIGRCETER 120  
 DB 61 DCERIGTSNYMAFPSPKALHARKHLEVEQSLSGSGQKASLOKSTIEKAKIGRCETER 120  
 QY 121 TRLAKELSLDQREQLKAEVEKYDCDPQVEEIRQANKVAKAANRWTDNIFAKSWA 180  
 DB 121 TRLAKELSLDQREQLKAEVEKYDCDPQVEEIRQANKVAKAANRWTDNIFAKSWA 180  
 QY 181 KRKFGFEENKIDRTFTGIPEDFDYID 205  
 DB 181 KRKFGFEENKIDRTFTGIPEDFDYID 205

RESULT 3  
 ADM83851  
 ID ADM83851 standard; protein; 205 AA.  
 AC ADM83851;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Human cancer gene 121P1 protein #6.  
 XX  
 KW Human; cancer gene 121P1; cytostatic; cancer; chromosome 4q; HLA;  
 KW human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;  
 KW colon cancer; lung cancer; pancreatic cancer; breast cancer;  
 KW cervical cancer; stomach cancer; gene therapy; vaccine.  
 OS Homo sapiens.  
 XX  
 PN US2003223997-A1.  
 XX  
 PD 04-DEC-2003.  
 XX  
 PF 28-FEB-2002; 2002US-00087190.  
 XX  
 PR 08-FEB-2001; 2001US-00779250.  
 XX

(CHAL/) CHALLITA-ETD P M.  
 PA (HUBE/) HUBERT R S.  
 PA (RAIT/) RAITANO A B.  
 PA (FARI/) FARIS M.  
 PA (APAR/) APAR D E H.  
 PA (GEW/) GE W.  
 PA (JAKO/) JAKOBOWITS A.

PI Chailita-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W,  
PI Jakobovits A;  
XX WPI; 2004-060522/06.  
XX  
PT New composition comprising a substance that modulates the status of  
PT 121PI1 or a molecule that is modulated by 121PI1, useful for detecting,  
PT treating or preventing cancer e.g. prostate, bladder, colon, breast or  
PT lung cancer.

Example 5; SEQ ID NO 61; 211pp; English.

XX The invention relates to a composition comprising a substance that  
XX modulates the status of 121PI1 (a protein encoded by a cancer expressed  
XX gene) or a molecule that is modulated by 121PI1 where status of the cell  
XX that expresses 121PI1 is modulated. Also included are a pharmaceutical  
XX composition comprising the novel composition in a human unit dose form, a  
XX recombinant protein comprising an antigen-binding region of a monoclonal  
XX antibody, a non-human transgenic animal that produces an antibody, a  
XX hybridoma that produces an antibody, a single chain monoclonal antibody  
XX that immunospecifically binds to a 121PI1-related protein (comprising  
XX the variable domains of the heavy and light chains of a monoclonal  
XX antibody), a vector comprising a polynucleotide that encodes a single  
XX chain monoclonal antibody, a polynucleotide that encodes an analogue  
XX peptide, delivering a cytotoxic agent or a diagnostic agent to a cell  
XX that expresses 121PI1, inhibiting growth of cancer cells that express  
XX 121PI1 (comprising administering to the cells the composition), treating  
XX a patient who bears cancer cells that express 121PI1, generating a  
XX mammalian immune response directed to 121PI1, inducing an immune  
XX response, monitoring 121PI1 gene products in a biological sample from a  
XX patient who has or who is suspected of having cancer, monitoring the  
XX presence of cancer in an individual and an assay for detecting the  
XX presence of a 121PI1-related protein or polynucleotide in a biological  
XX sample from a patient who has or who is suspected of having cancer. The  
XX composition may comprise a polynucleotide that comprises a 121PI1-  
XX related protein coding sequence provided that the coding sequence does  
XX not encode the entire amino acid sequence of 121PI1 (ADM83793). The  
XX substance also comprises a polynucleotide that encodes at least one  
XX peptide given in 16 Tables (given in the specification), the peptides  
XX being HLA (human leukocyte antigen)-binding epitopes from 121PI1 or its  
XX splice variants. The composition is useful for detecting, treating or  
XX preventing cancer, preferably prostate cancer, bladder cancer, kidney  
XX cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer,  
XX cervical cancer or stomach cancer. The composition can also be used as a  
XX vaccine to treat or prevent cancer that expresses or overexpresses  
XX 121PI1. The gene for 121PI1 is located on chromosome 4q. The present  
XX sequence is a 121PI1 protein (full-length or fragment).

XX Sequence 205 AA;

Query Match 100.0%; Score 1047; DB 8; Length 205;  
Best Local Similarity 100.0%; Pred. No. 1.5e-88;  
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSSEKRTMEIFSETKDVFOUKDEKTAPEKKGITVANSVEKVLDSIVDDGV 60  
DB 1 MSKKKGLSSEKRTMEIFSETKDVFOUKDEKTAPEKKGITVANSVEKVLDSIVDDGV 60  
QY 61 DCEKIGTSNYVAFPSKALHARKHKLVELESQLSSEGSQFASLOKSIIEFAKIGRCETEER 120  
DB 61 DCEKIGTSNYVAFPSKALHARKHKLVELESQLSSEGSQFASLOKSIIEFAKIGRCETEER 120  
QY 121 TRLAEELSLRQREOLKAVERKYDCDDQVVEEIRQANKVAKKANRWTNDINIFAKSWA 180  
DB 121 TRLAEELSLRQREOLKAVERKYDCDDQVVEEIRQANKVAKKANRWTNDINIFAKSWA 180  
QY 181 KRKFGFEENKIDRTFGIPEDFDYID 205  
DB 181 KRKFGFEENKIDRTFGIPEDFDYID 205

RESULT 4  
ADM83810

ID ADM83810 standard; protein; 205 AA.  
XX  
AC ADM83810;  
XX  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Human cancer gene 121PI1 protein #3.  
XX  
XX Human; cancer gene 121PI1; cytostatic; cancer; chromosome 4q; HLA;  
XX Human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;  
XX colon cancer; lung cancer; pancreatic cancer; breast cancer;  
XX cervical cancer; stomach cancer; gene therapy; vaccine.  
XX  
XX Homo sapiens.  
XX  
XX US2003223997-A1.  
XX  
XX 04-DEC-2003.  
XX  
XX 28-FEB-2002; 2002US-00087190.  
XX  
XX 08-FEB-2001; 2001US-00779250.  
XX  
XX (CHAL/) CHAILITA-EID P. M.  
XX (HUBE/) HUBERT R. S.  
XX (RAIT/) RAITANO A. B.  
XX (FARI/) FARIS M.  
XX (AFAR/) AFAR D E H.  
XX (GEW/) GE W.  
XX (JAKO/) JAKOBOVITS A.  
XX  
XX Chailita-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W,  
XX Jakobovits A;  
XX WPI; 2004-060522/06.  
XX  
XX New composition comprising a substance that modulates the status of  
XX 121PI1 or a molecule that is modulated by 121PI1, useful for detecting,  
XX treating or preventing cancer e.g. prostate, bladder, colon, breast or  
XX lung cancer.

PS Disclosure; SEQ ID NO 20; 211pp; English.

XX The invention relates to a composition comprising a substance that  
XX modulates the status of 121PI1 (a protein encoded by a cancer expressed  
XX gene) or a molecule that is modulated by 121PI1 where status of the cell  
XX that expresses 121PI1 is modulated. Also included are a pharmaceutical  
XX composition comprising the novel composition in a human unit dose form, a  
XX recombinant protein comprising an antigen-binding region of a monoclonal  
XX antibody, a non-human transgenic animal that produces an antibody, a  
XX hybridoma that produces an antibody, a single chain monoclonal antibody  
XX that immunospecifically binds to a 121PI1-related protein (comprising  
XX the variable domains of the heavy and light chains of a monoclonal  
XX antibody), a vector comprising a polynucleotide that encodes a single  
XX chain monoclonal antibody, a polynucleotide that encodes an analogue  
XX peptide, delivering a cytotoxic agent or a diagnostic agent to a cell  
XX that expresses 121PI1, inhibiting growth of cancer cells that express  
XX 121PI1 (comprising administering to the cells the composition), treating  
XX a patient who bears cancer cells that express 121PI1, generating a  
XX mammalian immune response directed to 121PI1, inducing an immune  
XX response, monitoring 121PI1 gene products in a biological sample from a  
XX patient who has or who is suspected of having cancer, monitoring the  
XX presence of cancer in an individual and an assay for detecting the  
XX presence of a 121PI1-related protein or polynucleotide in a biological  
XX sample from a patient who has or who is suspected of having cancer. The  
XX composition may comprise a polynucleotide that comprises a 121PI1-  
XX related protein coding sequence provided that the coding sequence does  
XX not encode the entire amino acid sequence of 121PI1 (ADM83793). The  
XX substance also comprises a polynucleotide that encodes at least one  
XX peptide given in 16 Tables (given in the specification), the peptides  
XX being HLA (human leukocyte antigen)-binding epitopes from 121PI1 or its  
XX splice variants. The composition is useful for detecting, treating or  
XX preventing cancer, preferably prostate cancer, bladder cancer, kidney

CC cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer,  
CC cervical cancer or stomach cancer. The composition can also be used as a  
CC vaccine to treat or prevent cancer. The composition can also be used as a  
CC 121p1. The gene for 121p1 is located on chromosome 4q. The present  
XX sequence is a 121p1 protein (full-length or fragment).  
SQ Sequence 205 AA;

Query Match  
Best Local Similarity 100.0%; Score 1047; DB 8; Length 205;  
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MSKKKGSAEKKTRMMEIFSETKDVOLKLEKIAPEKGTAMSVYVLSVDDGV 60  
QY 61 MSKKKGSAEKKTRMMEIFSETKDVOLKLEKIAPEKGTAMSVYVLSVDDGV 60  
DB 61 DCRIGTSNYWAFPSKALHARKHKLTVLESQSGSKASLQKSIKAKIGRCETEE 120  
QY 121 DCRIGTSNYWAFPSKALHARKHKLTVLESQSGSKASLQKSIKAKIGRCETEE 120  
DB 121 TRIAKELSLRDREOLKAEVEKYKDCPOVEEIRONANKVAKAANRWTDNIFAISMA 180  
QY 181 TRIAKELSLRDREOLKAEVEKYKDCPOVEEIRONANKVAKAANRWTDNIFAISMA 180  
DB 181 KRKGFEEKNKIDRTFGIPEDFDYID 205  
QY 205 KRKGFEEKNKIDRTFGIPEDFDYID 205

RESULT 5  
ADM83793  
ADM83793 standard; protein; 205 AA.

AC ADM83793;  
XX  
XX 03-JUN-2004 (first entry)  
DE Human cancer gene 121p1 protein #5.  
XX  
XX Human cancer gene 121p1; cytosolic; cancer; chromosome 4q; HLA;  
KW human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;  
KW cervical cancer; stomach cancer; gene therapy; vaccine.  
OS Homo sapiens.

XX US200323997-A1.  
XX 04-DEC-2003.  
XX  
XX 28-FEB-2002; 2002US-00087190.  
XX 08-FEB-2001; 2001US-00779250.  
XX  
XX (CHAL/) CHALLITA-EID P M.  
XX (HUBE/) HUBERT R S.  
XX (RAIT/) RAITANO A B.  
XX (FARI/) FARIS M.  
XX (AFAR/) AFAR D E H.  
XX (GEW/) GE W.  
XX (JAKO/) JAKOBOVITS A.  
XX  
XX Challa-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W,  
XX Jakovovits A;  
XX MPI; 2004-060522/06.  
XX  
XX New composition comprising a substance that modulates the status of  
XX 121p1 or a molecule that is modulated by 121p1, useful for detecting,  
XX creating or preventing cancer e.g. prostate, bladder, colon, breast or  
XX lung cancer.  
XX  
XX Disclosure; SEQ ID NO 22; 211p; English.

XX The invention relates to a composition comprising a substance that  
XX modulates the status of 121p1 (a protein encoded by a cancer gene  
XX gene) or a molecule that is modulated by 121p1 where status of  
XX that expresses 121p1 is modulated. Also included are a cancer expressed  
XX recombinant protein comprising the novel composition in a human unit dose form,  
XX antibody, a non-human transgenic animal that produces an antibody, a  
XX that immunospecifically binds to a 121p1-related protein (comprising  
XX the variable domain of the heavy and light chains of a monoclonal  
XX antibody), a vector comprising a polynucleotide that encodes a monoclonal  
XX chain monoclonal antibody, a polynucleotide that encodes a single  
XX peptide, delivering a cytotoxic agent or a diagnostic agent to a cell  
XX that expresses 121p1, inhibiting growth of cancer cells that expresses  
XX 121p1 (comprising administering to the cells the composition), treating  
XX a patient who bears cancer cells that express 121p1, generating a  
XX mammalian immune response directed to 121p1, inducing an immune  
XX response, monitoring 121p1 gene products in a biological sample from a  
XX patient who has or who is suspected of having cancer, monitoring the  
XX presence of cancer in an individual and an assay for detecting the  
XX sample from a patient who has or who is suspected of having cancer  
XX composition may comprise a polynucleotide that encodes a 121p1  
XX related protein coding sequence provided that the coding sequence does  
XX not encode the entire amino acid sequence of 121p1 (ADM83793). The  
XX substance also comprises a polynucleotide that encodes a 121p1  
XX peptide given in 16 Tables (given in the specification), the peptides  
XX being HLA (human leukocyte antigen)-binding epitopes from 121p1 or its  
XX splice variants. The composition is useful for detecting, treating or  
XX cancer, colon cancer, prostate cancer, pancreatic cancer, kidney  
XX cervical cancer or stomach cancer. The composition can also be used as a  
XX vaccine to treat or prevent cancer that expresses or overexpresses  
XX 121p1. The gene for 121p1 is located on chromosome 4q. The present  
XX sequence is a 121p1 protein (full-length or fragment).

Query Match  
Best Local Similarity 100.0%; Score 1047; DB 8; Length 205;  
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGSAEKKTRMMEIFSETKDVOLKLEKIAPEKGTAMSVYVLSVDDGV 60  
DB 1 MSKKKGSAEKKTRMMEIFSETKDVOLKLEKIAPEKGTAMSVYVLSVDDGV 60  
QY 61 DCRIGTSNYWAFPSKALHARKHKLTVLESQSGSKASLQKSIKAKIGRCETEE 120  
DB 61 DCRIGTSNYWAFPSKALHARKHKLTVLESQSGSKASLQKSIKAKIGRCETEE 120  
QY 121 TRIAKELSLRDREOLKAEVEKYKDCPOVEEIRONANKVAKAANRWTDNIFAISMA 180  
DB 121 TRIAKELSLRDREOLKAEVEKYKDCPOVEEIRONANKVAKAANRWTDNIFAISMA 180  
QY 181 KRKGFEEKNKIDRTFGIPEDFDYID 205  
DB 181 KRKGFEEKNKIDRTFGIPEDFDYID 205

RESULT 6  
ADM83793  
ADM83793 standard; protein; 205 AA.

AC ADM83793;  
XX  
XX 03-JUN-2004 (first entry)  
DE Human cancer gene 121p1 protein #1.  
XX  
XX Human cancer gene 121p1; cytosolic; cancer; chromosome 4q; HLA;  
KW human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;  
KW cervical cancer; stomach cancer; gene therapy; vaccine.  
OS Homo sapiens.

KM	cervical cancer; stomach cancer; gene therapy; vaccine.
XX	
OS	Homo sapiens.
XX	
PN	US2003223997-A1.
XX	
PD	04-DEC-2003.
XX	
PF	28-FEB-2002; 2002US-00087190.
XX	
PR	08-FEB-2001; 2001US-00779250.
XX	
PA	(CHAL/) CHALLITA-EID P M.
PA	(HUBE/) HUBERT R S.
PA	(RAIT/) RAITANO A B.
PA	(FARI/) FARIS W.
PA	(AFAR/) AFAR D E H.
PA	(GEWV/) GE W.
PA	(JAKO/) JAKOBOVITS A.
XX	
PI	Challita-Eid PM, Hubert RS, Raitano AB, Faris W, Afar DEH, Ge W, Jakobovits A;
XX	
DR	WPI: 2004-060522/06.
DR	N-PSSB; ADM83792.
XX	
PT	New composition comprising a substance that modulates the status of 12lPIFI or a molecule that is modulated by 12lPIFI, useful for detecting, treating or preventing cancer e.g. prostate, bladder, colon, breast or lung cancer.
PT	
DS	Claim 19; SEQ ID NO 3; 21lpp; English.

The invention relates to a composition comprising a substance that modulates the status of 121P1F1 (a protein encoded by a cancer expressed gene) or a molecule that is modulated by 121P1F1 where status of the cell that expresses 121P1F1 is modulated. Also included are a pharmaceutical composition comprising the novel composition in a human unit dose form, a recombinant protein comprising an antigen-binding region of a monoclonal antibody, a non-human transgenic animal that produces an antibody, a hybridoma that produces an antibody, a single chain monoclonal antibody that immunospecifically binds to a 121P1F1-related protein (comprising the variable domains of the heavy and light chains of a monoclonal antibody), a vector comprising a polynucleotide that encodes a single chain monoclonal antibody, a polynucleotide that encodes an analogue peptide, delivering a cytotoxic agent or a diagnostic agent to a cell that expresses 121P1F1, inhibiting growth of cancer cells that express 121P1F1 (comprising administering to the cells the composition), treating a patient who bears cancer cells that express 121P1F1, generating a mammalian immune response directed to 121P1F1, inducing an immune response, monitoring 121P1F1 gene products in a biological sample from a patient who has or who is suspected of having cancer, monitoring the presence of cancer in an individual and an assay for detecting the presence of a 121P1F1-related protein or polynucleotide in a biological sample from a patient who has or who is suspected of having cancer. The composition may comprise a polynucleotide that comprises a 121P1F1-related protein coding sequence provided that the coding sequence does not encode the entire amino acid sequence of 121P1F1 (ADM83793). The substance also comprises a polynucleotide that encodes at least one peptide given in 16 Tables (given in the specification), the peptides being HLA (human leukocyte antigen)-binding epitopes from 121P1F1 or its splice variants. The composition is useful for detecting, treating or preventing cancer, preferably prostate cancer, bladder cancer, kidney cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer, cervical cancer or stomach cancer. The composition can also be used as a vaccine to treat or prevent cancer that expresses or overexpresses 121P1F1. The gene for 121P1F1 is located on chromosome 4q. The present sequence is a 121P1F1 protein (full-length or fragment).

XX	Sequence 205 AA;
Query Match	100.0%; Score 1047; DB 8; Length 205;
Beet Local Similarity	100.0%; Pred. No. 1,5e-88;

	Matches	205;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0		
Qy	1	MSKKKG	SAEKKRR	RRMWEI	FSETKD	VPOLK	OLEKIA	PKKXGTTAS	VKEVLS	LOSLVDGAV	60	
Db	1	MSKKKG	SAEKKRR	RRMWEI	FSETKD	VPOLK	OLEKIA	PKKXGTTAS	VKEVLS	LOSLVDGAV	60	
Qy	61	DCEKIG	TSNYWAF	PPSKAL	HARKKL	EVLES	QSLSGS	QKHA	SLQSI	IEKAKIG	RCETEE	120
Db	61	DCEKIG	TSNYWAF	PPSKAL	HARKKL	EVLES	QSLSGS	QKHA	SLQSI	IEKAKIG	RCETEE	120
Qy	121	TRLAKEL	SSLDPQ	SEOLKAE	VEKTKY	DCDP	OVVEE	ELRQ	NKAKY	KEAAN	RTWDNI	180
Db	121	TRLAKEL	SSLDPQ	SEOLKAE	VEKTKY	DCDP	OVVEE	ELRQ	NKAKY	KEAAN	RTWDNI	180
Qy	181	KRKGF	PEENK	IDRTFG	IPED	FYDID	205					
Db	181	KRKGF	PEENK	IDRTFG	IPED	FYDID	205					

CC	that immunospecifically binds to a 121p1F1-related protein (comprising
CC	antibody, a non-human transgenic animal that produces an antibody, a
CC	recombinant protein comprising an antigen-binding region of a monoclonal
CC	composition comprising the novel composition in a human unit dose form, a
CC	gene) or a molecule that is modulated by 121p1F1 where status of the cell
CC	modulates the status of 121p1F1 (a protein encoded by a cancer expressed
CC	The invention relates to a composition comprising a substance that
XX	disclosure; SEQ ID NO 21; 211pp; English.
PT	lung cancer.
PT	treating or preventing cancer e.g. prostate, bladder, colon, breast or
XX	121p1F1 or a molecule that is modulated by 121p1F1, useful for detecting,
XX	New composition comprising a substance that modulates the status of
XX	WPI; 2004-060522/06.
PI	Jakobovits A;
PI	Challita-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;
XX	(JAKO/) JAKOBOVITS A.
PA	(GEWM/) GE W.
PA	(AFAR/) AFAR D E H.
PA	(FARI/) FARIS M.
PA	(RAIT/) RAITANO A B.
PA	(HUBE/) HUBERT R S.
XX	(CHAL/) CHALLITA-EID P M.
XX	08-FEB-2001; 2001US-00779250.
XX	28-FEB-2002; 2002US-00087190.
XX	04-DEC-2003.
XX	US2003223997-A1.
XX	Homo sapiens.
XX	cervical cancer; stomach cancer; gene therapy; vaccine.
KW	colon cancer; lung cancer; pancreatic cancer; breast cancer;
KW	human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;
KW	Human cancer gene 121p1F1 protein #4.
DE	Human cancer gene 121p1F1 protein #4.
XX	03-JUN-2004 (first entry)
XX	ADM83811;
XX	ADM83811 standard; protein; 205 AA.
ID	ADM83811
RESULT 7	



DB 121 TRLAKELSLRQREQLKAEVEKYDCDPQVVEIRQANKVAKKANRWTDNIFAIKSWA 180  
 QY 181 KRKFGFEENKIDRTFGIPEDFDYID 205  
 DB 181 KRKFGFEENKIDRTFGIPEDFDYID 205

RESULT 9  
 ABP75541  
 ID ABP75541 standard; protein; 219 AA.  
 AC ABP75541;  
 DT 10-FEB-2003 (first entry)  
 XX  
 DE Human secretory polypeptide SPTM SEQ ID NO 725.  
 XX  
 KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;  
 KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;  
 KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;  
 KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;  
 KW anti-inflammatory; immunosuppressive; neuroprotective; nocotropic;  
 KW neuroleptic; anticonvulsant; cytosolic; antiparkinsonian; anxiolytic;  
 KW antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;  
 KW secretory polynucleotide; secretory protein.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283876-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 27-MAR-2002; 2002WO-US009921.  
 XX  
 PR 29-MAR-2001; 2001US-0280067P.  
 PR 29-MAR-2001; 2001US-0280068P.  
 PR 16-MAY-2001; 2001US-0291280P.  
 PR 17-MAY-2001; 2001US-0291829P.  
 PR 17-MAY-2001; 2001US-0291849P.  
 PR 19-JUN-2001; 2001US-0299428P.  
 PR 20-JUN-2001; 2001US-0299776P.  
 PR 20-JUN-2001; 2001US-0300001P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
 PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;  
 PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstlin EH;  
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;  
 XX  
 DR WPI; 2003-075543/07.  
 DR N-PSDB; ABZ35987.  
 XX  
 PT New human secretory proteins and polynucleotides, useful for diagnosing,  
 PT treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),  
 PT neurological disorders (e.g. Alzheimer's), or cell proliferations or  
 PT cancers.  
 XX  
 PS Claim 27; SEQ ID NO 725; 458bp + Sequence Listing; English.  
 XX  
 CC The invention relates to a secretory polynucleotide (designated spm)  
 CC comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a  
 CC naturally occurring polynucleotide sequence at least 90 % identical to  
 CC the polynucleotide sequence, a polynucleotide complementary to them or an  
 CC RNA equivalent of them. The polypeptide or polynucleotide are useful for  
 CC treating, preventing or diagnosing a disease or condition associated with  
 CC the expression of functional SPTM. These are particularly useful for  
 CC diagnosing, treating or preventing autoimmune/inflammatory disorders  
 CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's  
 CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,  
 CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,

CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,  
 CC schizophrenia or amnesia), or cell proliferative disorders (e.g.  
 CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,  
 CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,  
 CC breast, cervix or prostate). The present sequence is one of the SPTM  
 CC proteins of the invention (ABP75384-ABP75962). Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 219 AA;  
 Query Match 100.0%; Score 1047; DB 6; Length 219;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-88;  
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSNEKRTMMEIFSETKVPQKDKLEKAPKPKGTITANSYVETLSVDDGV 60  
 DB 15 MSKKKGLSNEKRTMMEIFSETKVPQKDKLEKAPKPKGTITANSYVETLSVDDGV 74  
 QY 61 DDERIGTSNYWAFPSKALHARKHLEVLSEQLSGSQGHASLOKSIERAKIGRCETEE 120  
 DB 75 DDERIGTSNYWAFPSKALHARKHLEVLSEQLSGSQGHASLOKSIERAKIGRCETEE 134  
 QY 121 TRLAKELSLRQREQLKAEVEKYDCDPQVVEIRQANKVAKKANRWTDNIFAIKSWA 180  
 DB 135 TRLAKELSLRQREQLKAEVEKYDCDPQVVEIRQANKVAKKANRWTDNIFAIKSWA 194

QY 181 KRKFGFEENKIDRTFGIPEDFDYID 205  
 DB 195 KRKFGFEENKIDRTFGIPEDFDYID 219

RESULT 10  
 ADM83834  
 ID ADM83834 standard; protein; 206 AA.  
 XX  
 AC ADM83834;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Human cancer gene 121P1F1 variant protein #1.  
 XX  
 KW Human; cancer gene 121P1F1; cytosolic; cancer; chromosome 4q; HLA;  
 KW human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;  
 KW colon cancer; lung cancer; pancreatic cancer; breast cancer;  
 KW cervical cancer; stomach cancer; gene therapy; vaccine.  
 OS Homo sapiens.  
 XX  
 PN US2003223997-A1.  
 XX  
 PD 04-DEC-2003.  
 XX  
 PF 28-FEB-2002; 2002US-00087190.  
 XX  
 PR 08-FEB-2001; 2001US-00779250.  
 XX  
 PA (CHAL/) CHALLITA-EID P M.  
 PA (HUBE/) HUBERT R S.  
 PA (RAIT/) RAITANO A B.  
 PA (FARI/) FARIS M.  
 PA (AFAR/) AFAR D E H.  
 PA (GEWW/) GE W.  
 PA (JAKO/) JAKOBOVITS A.  
 PI Challice-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;  
 PI Jakobovits A;  
 XX  
 DR WPI; 2004-060522/06.  
 XX  
 PT New composition comprising a substance that modulates the status of  
 PT 121P1F1 or a molecule that is modulated by 121P1F1, useful for detecting,



invention relates to a composition comprising a substance that modulates the status of 121p1 (a protein encoded by a cancer gene) or a molecule that is modulated by 121p1 where status of the cell that expresses 121p1 is modulated. Also included are the cell recombinant protein comprising the novel composition in a human unit dose form, a hybridoma that produces an antigen-binding region of a monoclonal antibody, a non-human transgenic animal that produces an antibody, a cell that immunospecifically binds to a 121p1-related protein (comprising the variable domains of the heavy and light chains of a monoclonal antibody), a vector comprising a polynucleotide that encodes a single chain monoclonal antibody, a polynucleotide that encodes an analogue peptide, delivering a cytotoxic agent or a diagnostic agent to a cell 121p1 (comprising administering to the cells the composition), treating a patient who bears cancer cells that express 121p1, generating a mammalian immune response directed to 121p1, inducing an immune response, monitoring 121p1 gene products in a biological sample from a patient who has or who is suspected of having cancer, monitoring the presence of cancer in an individual and an assay for detecting the presence of a 121p1-related protein or polynucleotide in a biological sample from a patient who has or who is suspected of having cancer. The composition may comprise a polynucleotide that comprises a 121p1-related protein coding sequence provided that the coding sequence does not encode the entire amino acid sequence of 121p1 (ADW83793). The substance also comprises a polynucleotide that encodes at least one peptide given in 16 Tables (given in the specification), the peptides being HLA (human leukocyte antigen)-binding epitopes from 121p1 or its splice variants. The composition is useful for detecting, treating or preventing cancer, preferably prostate cancer, bladder cancer, kidney cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer, cervical cancer or stomach cancer. The composition can also be used as vaccine to treat or prevent cancer that expresses or overexpresses 121p1. The gene for 121p1 is located on chromosome 4q. The present sequence is a 121p1 protein (full-length or fragment).

	0	Mismatches	0	Indels	1	Gaps	1
1	MSKKKGLSAEKRT	RMWEIFSE	FDVFO	KDLEKIAPEKGI	TAMSYKEV	LSI	VDGDMV
1	MSKKKGLSAEKRT	RMWEIFSE	KDVFOLKE	IAPEKGI	TAMSYKEV	LSI	VDGDMV
61	DCERIGTSNY	YMAFP	PSKALHAR	KHL	EVL	ES	-QLSEGSQHASL
61	DCERIGTSNY	YMAFP	PSKALHAR	KHL	EVL	ESQ	LSSEGSQHASL
61	DCERIGTSNY	YMAFP	PSKALHAR	KHL	EVL	ESQ	LSSEGSQHASL
120	RTRLAKEL	SSL	RSOR	QOL	KAVEY	KKOD	QVVEEIR
120	RTRLAKEL	SSL	RSOR	QOL	KAVEY	KKOD	QVVEEIR
121	RTRLAKEL	SSL	RSOR	QOL	KAVEY	KKOD	QVVEEIR
180	AKKGFEE	NNKID	RTFGI	PE	DFY	ID	205
181	AKKGFEE	NNKID	RTFGI	PE	DFY	ID	206

XX Human, cancer gene 121P1; cytostatic; cancer; chromosome 4q; HLA,  
KW human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer,  
KW colon cancer; lung cancer; pancreatic cancer; breast cancer;  
KW cervical cancer; stomach cancer; gene therapy; vaccine.  
OS Homo sapiens.  
XX  
XX US200323997-A1.  
XX  
XX 04-DEC-2003.  
XX  
XX 28-FEB-2002; 2002US-00087190.  
XX  
XX 08-FEB-2001; 2001US-00779250.  
XX  
XX (CHAL/) CHALLITA-EID P M.  
XX (HUBE/) HUBERT R S.  
XX (RAIT/) RAITANO A B.  
XX (FARI/) FARIS M E.  
XX (AFAR/) AFAR D E H.  
XX (GEWM/) GE W.  
XX (JAKO/) JAKOVBITS A.  
XX  
XX Challita-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W,  
XX Jakobovits A;  
XX WPI: 2004-060522/06.  
XX  
XX New composition comprising a substance that modulates the status of  
XX 121P1 or a molecule that is modulated by 121P1 where status of the cell  
XX treating or preventing cancer e.g. prostate, bladder, colon, breast or  
XX lung cancer.  
XX  
XX Example 5; SEQ ID NO 45; 211pp; English.  
XX  
XX The invention relates to a composition comprising a substance that  
XX modulates the status of 121P1 (a protein encoded by a cancer expressed  
XX gene) or a molecule that is modulated by 121P1 where status of the cell  
XX that expresses 121P1 is modulated. Also included are a pharmaceutical  
XX composition comprising the novel composition in a human unit dose form,  
XX recombinant protein comprising an antigen-binding region of a monoclonal  
XX antibody, a non-human transgenic animal that produces an antibody, a  
XX hybridoma that produces an antibody, a single chain monoclonal antibody  
XX that immunospecifically binds to a 121P1-related protein (comprising  
XX the variable domains of the heavy and light chains of a monoclonal  
XX antibody), a vector comprising a polynucleotide that encodes a single  
XX chain monoclonal antibody, a polynucleotide that encodes an analogue  
XX peptide, delivering a cytotoxic agent or a diagnostic agent to a cell  
XX that expresses 121P1, inhibiting growth of cancer cells that express  
XX 121P1 (comprising administering to the cells the composition), treating  
XX mammalian immune response directed to 121P1, generating an immune  
XX response, monitoring 121P1 gene products in a biological sample from a  
XX patient who has or who is suspected of having cancer, monitoring the  
XX presence of cancer in an individual and an assay for detecting the  
XX presence of a 121P1-related protein or polynucleotide in a biological  
XX sample from a patient who has or who is suspected of having cancer. The  
XX composition may comprise a polynucleotide that comprises a 121P1-  
XX related protein coding sequence provided that the coding sequence does  
XX not encode the entire amino acid sequence of 121P1 (NDM83793). The  
XX substance also comprises a polynucleotide that encodes at least one  
XX being HLA (human leukocyte antigen)-binding epitopes from 121P1 or its  
XX splice variants. The composition is useful for detecting, treating or  
XX preventing cancer. The composition is useful for detecting, treating or  
XX cervical cancer, lung cancer, pancreatic cancer, bladder cancer, kidney  
XX cancer, colon cancer, stomach cancer, pancreatic cancer, breast cancer,  
XX vaccine to treat or prevent cancer. The composition can also be used as a  
XX 121P1. The gene for 121P1 is located on chromosome 4q. The present  
XX sequence is a 121P1 protein (full-length or fragment).  
XX  
XX Sequence 206 AA:  
XX



Query Match 99.0%; Score 1036.5; DB 8; Length 206;  
 Best Local Similarity 99.5%; Pred. No. 1.4e-87;  
 Matches 205; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MSKKGLSAEKKRTMMEIFSETKDVFOJKDLKIAPKKEGTTAMSKVEVLQSLVDGMY 60  
 DB 1 MSKKGLSAEKKRTMMEIFSETKDVFOJKDLKIAPKKEGTTAMSKVEVLQSLVDGMY 60  
 QY 61 DCEKGTGNYWAFPSKALHARKHLEVESQ-LSRGSQKHAISLOKSIKAKTGRCTEE 119  
 DB 61 DCEKGTGNYWAFPSKALHARKHLEVESQ-LSRGSQKHAISLOKSIKAKTGRCTEE 120  
 QY 120 RTRLAKELSLRDQREQLKAEVEKYKCDPQVVEEIRQANKYAKKANRMTDNIFAISK 179  
 DB 121 RTRLAKELSLRDQREQLKAEVEKYKCDPQVVEEIRQANKYAKKANRMTDNIFAISK 180  
 QY 180 AKRKGFEENKIDRTFGIPEDFDYID 205  
 DB 181 AKRKGFEENKIDRTFGIPEDFDYID 206

RESULT 12  
 ADM83814  
 ID ADM83814 standard; protein; 198 AA.  
 AC ADM83814;  
 DT 03-JUN-2004 (first entry)  
 XX Human cancer gene 121P1F1 protein fragment.  
 DE  
 XX  
 KW Human; cancer gene 121P1F1; cytostatic; cancer; chromosome 4q; HLA;  
 KW human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;  
 KW colon cancer; lung cancer; pancreatic cancer; breast cancer;  
 KW cervical cancer; stomach cancer; gene therapy; vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003223997-A1.  
 XX  
 PD 04-DEC-2003.  
 XX  
 PF 28-FEB-2002; 2002US-00087190.  
 PR 08-FEB-2001; 2001US-00779250.  
 XX  
 PA (CHAL/) CHALLITA-EID P. M.  
 PA (HUBE/) HUBERT R. S.  
 PA (RAIT/) RAITANO A. B.  
 PA (FARI/) FARIS M.  
 PA (AFAR/) AFAR D. E. H.  
 PA (GEWV/) GE W.  
 PA (JAKO/) JAKOBOVITS A.  
 XX  
 PI Challita-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;  
 PI Jakobovits A;  
 DR WPI; 2004-060522/06.  
 XX  
 PT New composition comprising a substance that modulates the status of  
 PT 121P1F1 or a molecule that is modulated by 121P1F1, useful for detecting,  
 PT treating or preventing cancer e.g. prostate, bladder, colon, breast or  
 PT lung cancer.  
 XX  
 PS Example 2; SEQ ID NO 24; 211pp; English.  
 XX  
 CC The invention relates to a composition comprising a substance that  
 CC modulates the status of 121P1F1 (a protein encoded by a cancer expressed  
 CC gene) or a molecule that is modulated by 121P1F1 where status of the cell  
 CC that expresses 121P1F1 is modulated. Also included are a pharmaceutical  
 CC composition comprising the novel composition in a human unit dose form, a  
 CC recombinant protein comprising an antigen-binding region of a monoclonal

CC antibody, a non-human transgenic animal that produces an antibody, a  
 CC hybridoma that produces an antibody, a single chain monoclonal antibody  
 CC that immunospecifically binds to a 121P1F1-related protein (comprising  
 CC the variable domains of the heavy and light chains of a monoclonal  
 CC antibody), a vector comprising a polynucleotide that encodes a single  
 CC chain monoclonal antibody, a polynucleotide that encodes an analogue  
 CC peptide, delivering a cytotoxic agent or a diagnostic agent to a cell  
 CC that expresses 121P1F1, inhibiting growth of cancer cells that expresses  
 CC 121P1F1 (comprising administering to the cells the composition), treating  
 CC a patient who bears cancer cells that express 121P1F1, generating a  
 CC mammalian immune response directed to 121P1F1, inducing an immune  
 CC response, monitoring 121P1F1 gene products in a biological sample from a  
 CC patient who has or who is suspected of having cancer, monitoring the  
 CC presence of cancer in an individual and an assay for detecting the  
 CC presence of a 121P1F1-related protein or polynucleotide in a biological  
 CC sample from a patient who has or who is suspected of having cancer. The  
 CC composition may comprise a polynucleotide that comprises a 121P1F1-  
 CC related protein coding sequence provided that the coding sequence does  
 CC not encode the entire amino acid sequence of 121P1F1 (ADM83793). The  
 CC substance also comprises a polynucleotide that encodes at least one  
 CC peptide given in 16 Tables (given in the specification), the peptides  
 CC being HLA (human leukocyte antigen)-binding epitopes from 121P1F1 or its  
 CC splice variants. The composition is useful for detecting, treating or  
 CC preventing cancer, preferably prostate cancer, bladder cancer, kidney  
 CC cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer,  
 CC cervical cancer or stomach cancer. The composition can also be used as a  
 CC vaccine to treat or prevent cancer that expresses or overexpresses  
 CC 121P1F1. The gene for 121P1F1 is located on chromosome 4q. The present  
 CC sequence is a 121P1F1 protein (full-length or fragment).  
 XX  
 SQ Sequence 198 AA;

Query Match 96.6%; Score 1011; DB 8; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-85;

Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KGLSAEKKRTMMEIFSETKDVFOJKDLKIAPKKEGTTAMSKVEVLQSLVDGMYDCR 64  
 DB 1 KGLSAEKKRTMMEIFSETKDVFOJKDLKIAPKKEGTTAMSKVEVLQSLVDGMYDCR 60  
 QY 65 IGTSNYYWAFPSKALHARKHLEVESQ-LSRGSQKHAISLOKSIKAKTGRCTEERTLA 124  
 DB 61 IGTSNYYWAFPSKALHARKHLEVESQ-LSRGSQKHAISLOKSIKAKTGRCTEERTLA 120  
 QY 125 KELSLRDQREQLKAEVEKYKCDPQVVEEIRQANKYAKKANRMTDNIFAISKAKRF 184  
 DB 121 KELSLRDQREQLKAEVEKYKCDPQVVEEIRQANKYAKKANRMTDNIFAISKAKRF 180  
 QY 185 GFENKIDRTFGIPEDFD 202  
 DB 181 GFENKIDRTFGIPEDFD 198

RESULT 13  
 AAM40043  
 ID AAM40043 standard; protein; 190 AA.  
 AC AAM40043;  
 DT 22-OCT-2001 (first entry)  
 XX Human polypeptide SEQ ID NO 3188.  
 DE  
 XX  
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokine;  
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.

XX 26-JUL-2001.  
 XX 26-DEC-2000; 2000MO-US034263.  
 PF 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-00552317.  
 PR 20-JUN-2000; 2000US-00552317.  
 PR 19-JUL-2000; 2000US-00620312.  
 PR 03-AUG-2000; 2000US-00653450.  
 PR 14-SEP-2000; 2000US-00662191.  
 PR 19-OCT-2000; 2000US-00693036.  
 PR 29-NOV-2000; 2000US-00727344.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
 PR Zhou P, Goodrich R, Drmanac RT;  
 DR WPI: 2001-442253/47.  
 DR N-PSDB; AA159199.  
 XX Novel nucleic acids and polypeptides, useful for treating disorders such  
 PT as central nervous system injuries.  
 XX Example 4; SEQ ID NO 3188; 10078bp; English.  
 PS The invention relates to human nucleic acids (AA157798-AA161369) and the  
 CC encoded polypeptides (AA158642-AA162213) with neurotrophic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localized neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilization of the activities such as: immune system suppression,  
 CC activation/inhibition of activities, chemotactic/chemokinetic activity,  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukemias and  
 CC C.N.S disorders. Note: The sequence data for this patent did not form  
 CC part of the printed specification  
 XX Sequence 190 AA;  
 SQ  
 Query Match 93.1%; Score 975; DB 4; Length 190;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-82;  
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 16 MMEIFSEKDVFOLEKIAPEKIGITAMSVKYLQSLVDDGWDVCERIGTSNYVAAP 75  
 DB 1 MMEIFSEKDVFOLEKIAPEKIGITAMSVKYLQSLVDDGWDVCERIGTSNYVAAP 75  
 QY 76 SKALHARKHLEVLSESGOKASLOKSTIEKAKIGRCETERTRLAKELSLDDORE 135  
 DB 61 SKALHARKHLEVLSESGOKASLOKSTIEKAKIGRCETERTRLAKELSLDDORE 135  
 QY 136 OLKAEVEKTKDDCPQVEIRQANKVAKAANRWTDNIFAIKSMARKKGFENKIDRTF 195  
 DB 121 OLKAEVEKTKDDCPQVEIRQANKVAKAANRWTDNIFAIKSMARKKGFENKIDRTF 195  
 QY 196 GIPEDFDYID 205  
 DB 181 GIPEDFDYID 190

XX 29-JAN-2004 (first entry)  
 DT 121PIPI variant 4 protein.  
 DE 121PIPI, 121PIPI modulation; human; chromosome 4q; cytostatic;  
 KW gene therapy; vaccine; cancer; immune response; immunisation.  
 XX OS Synthetic.  
 XX Homo sapiens.  
 XX WO200295009-A2.  
 PD 28-NOV-2002.  
 XX 28-FEB-2002; 2002MO-US006242.  
 PF 05-MAR-2001; 2001US-00799250.  
 PR (AGEN-) AGENSYS INC.  
 PA Challita-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DH, Ge W;  
 PI Jakobovits AJ;  
 DR WPI: 2003-156757/15.  
 DR N-PSDB; ADD84546.  
 XX Composition comprising a substance that modulates the status of 121PIPI,  
 PT useful in diagnosing, preventing, prognosticating or treating patients  
 PT with cancer that expresses 121PIPI, such as breast, colon, ovarian or  
 PT lung cancer.  
 PS Claim 19; Fig 2F; 285bp; English.  
 CC The present invention describes a composition (I) comprising a substance  
 CC that modulates the status of 121PIPI (gene and encoded protein), or a  
 CC molecule that is modulated by 121PIPI, where the status of a cell that  
 CC 4q. (I) has cytostatic activity, and can be used in gene therapy, and in  
 CC vaccines. The composition (I) can be used for diagnosing, preventing,  
 CC prognosticating or treating patients with cancer that expresses 121PIPI,  
 CC such as breast, colon, ovarian or lung cancer. The 121PIPI gene or its  
 CC fragment can be used to elicit a humoral or cellular immune response.  
 CC 121PIPI antibodies can be used in active or passive immunisation. 121PIPI  
 CC polynucleotides are useful as probes and primers for the amplification or  
 CC expression of 121PIPI genes, as coding sequences for directing the  
 CC inhibiting the expression of 121PIPI genes, or as tools for modulating or  
 CC in the exemplification of the present invention.  
 XX Sequence 190 AA;  
 SQ  
 Query Match 93.1%; Score 975; DB 7; Length 190;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-82;  
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 16 MMEIFSEKDVFOLEKIAPEKIGITAMSVKYLQSLVDDGWDVCERIGTSNYVAAP 75  
 DB 1 MMEIFSEKDVFOLEKIAPEKIGITAMSVKYLQSLVDDGWDVCERIGTSNYVAAP 75  
 QY 76 SKALHARKHLEVLSESGOKASLOKSTIEKAKIGRCETERTRLAKELSLDDORE 135  
 DB 61 SKALHARKHLEVLSESGOKASLOKSTIEKAKIGRCETERTRLAKELSLDDORE 135  
 QY 136 OLKAEVEKTKDDCPQVEIRQANKVAKAANRWTDNIFAIKSMARKKGFENKIDRTF 195  
 DB 121 OLKAEVEKTKDDCPQVEIRQANKVAKAANRWTDNIFAIKSMARKKGFENKIDRTF 195  
 QY 196 GIPEDFDYID 205  
 DB 181 GIPEDFDYID 190

RESULT 15  
ADM83857  
ADM83857 standard; protein; 190 AA.  
XX  
AC ADM83857;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Human cancer gene 121P1F1 protein 16-205.  
XX  
KM Human; cancer gene 121P1F1; cytosolic; cancer; chromosome 4q; HLA;  
KM human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;  
KM colon cancer; lung cancer; pancreatic cancer; breast cancer;  
KM cervical cancer; stomach cancer; gene therapy; vaccine.  
XX  
OS Homo sapiens.  
XX  
FN US2003223997-A1.  
XX  
PD 04-DEC-2003.  
XX  
PF 28-FEB-2002; 2002US-00087190.  
PR 08-FEB-2001; 2001US-0079250.  
XX  
PA (CHAL/) CHALLITA-BID P M.  
PA (HUBE/) HUBERT R S.  
PA (RAT/) RATANO A B.  
PA (FARI/) FARIS M.  
PA (AFAR/) AFAR D E H.  
PA (GEWU/) GE W.  
PA (JAKO/) JAKOBOVITS A.  
XX  
PI Challaite-Bid PM, Hubert RS, Raicano AB, Faris M, Afar DEH, Ge W;  
PI Jakobovits A;  
XX  
DR WPI; 2004-060522/06.  
XX  
PT New composition comprising a substance that modulates the status of  
PT 121P1F1 or a molecule that is modulated by 121P1F1, useful for detecting,  
PT treating or preventing cancer e.g. prostate, bladder, colon, breast or  
PT lung cancer.  
XX  
PS Example 5; SEQ ID NO 67; 211pp; English.  
XX  
CC The invention relates to a composition comprising a substance that  
CC modulates the status of 121P1F1 (a protein encoded by a cancer expressed  
CC gene) or a molecule that is modulated by 121P1F1 where status of the cell  
CC that expresses 121P1F1 is modulated. Also included are a pharmaceutical  
CC composition comprising the novel composition in a human unit dose form, a  
CC recombinant protein comprising an antigen-binding region of a monoclonal  
CC antibody, a non-human transgenic animal that produces an antibody, a  
CC hybridoma that produces an antibody, a single chain monoclonal antibody  
CC that immunospecifically binds to a 121P1F1-related protein (comprising  
CC the variable domains of the heavy and light chains of a monoclonal  
CC antibody), a vector comprising a polynucleotide that encodes a single  
CC chain monoclonal antibody, a polynucleotide that encodes an analogue  
CC peptide, delivering a cytotoxic agent or a diagnostic agent to a cell  
CC that expresses 121P1F1, inhibiting growth of cancer cells that expresses  
CC 121P1F1 (comprising administering to the cells the composition), treating  
CC a patient who bears cancer cells that express 121P1F1, generating a  
CC mammalian immune response directed to 121P1F1, inducing an immune  
CC response, monitoring 121P1F1 gene products in a biological sample from a  
CC patient who has or who is suspected of having cancer, monitoring the  
CC presence of cancer in an individual and an assay for detecting the  
CC presence of a 121P1F1-related protein or polynucleotide in a biological  
CC sample from a patient who has or who is suspected of having cancer. The  
CC composition may comprise a polynucleotide that comprises a 121P1F1-  
CC related protein coding sequence provided that the coding sequence does  
CC not encode the entire amino acid sequence of 121P1F1 (ADM83793). The  
CC substance also comprises a polynucleotide that encodes at least one  
CC peptide given in 16 tables (given in the specification), the peptides  
CC being HLA (human leukocyte antigen)-binding epitopes from 121P1F1 or its

CC splice variants. The composition is useful for detecting, treating or  
CC preventing cancer, preferably prostate cancer, bladder cancer, kidney  
CC cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer,  
CC cervical cancer or stomach cancer. The composition can also be used as a  
CC vaccine to treat or prevent cancer that expresses or overexpresses  
CC 121P1F1. The gene for 121P1F1 is located on chromosome 4q. The present  
CC sequence is a 121P1F1 protein (full-length or fragment).  
XX  
SQ Sequence 190 AA;  
Query Match 93.1%; Score 975; DB 8; Length 190;  
Best Local Similarity 100.0%; Pred. No. 6.3e-82;  
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 16 MMEIFSETKDVLPQDLKELIAPKEKGIPTAMSVKVELQSLVDGMDCEKIGTSNYMARP 75  
DB 1 MMEIFSETKDVLPQDLKELIAPKEKGIPTAMSVKVELQSLVDGMDCEKIGTSNYMARP 60  
QY 76 SKALHARKKLELVLESQSEGSQKHAISLOKSTIEKAKIGRCETERTRLAKELSLRDQRE 135  
DB 61 SKALHARKKLELVLESQSEGSQKHAISLOKSTIEKAKIGRCETERTRLAKELSLRDQRE 120  
QY 136 QLKAEVEKKDCDPQVEVERIRQANKVAKAARWTDNIFAIKSNKRRKGFENKIDRF 195  
DB 121 QLKAEVEKKDCDPQVEVERIRQANKVAKAARWTDNIFAIKSNKRRKGFENKIDRF 180  
QY 196 GIPEDPDYID 205  
DB 181 GIPEDPDYID 190

Search completed: March 23, 2005, 19:14:48  
Job time : 171 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2005, 19:09:32 ; Search time 43 Seconds  
(without alignments)  
355.885 Million cell updates/sec

Title: US-10-087-190-3

Perfect score: 1047  
Sequence: 1 MSKKGSLAEKRTMMEIF.....FRENKIDRTFGIPEDFYID 205

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	593.5	56.7	127	US-09-621-976-4959	Sequence 4959, Ap
2	191.5	18.3	260	US-09-248-796A-14485	Sequence 14485, A
3	134.5	12.8	1270	US-09-538-092-1321	Sequence 1321, Ap
4	117	11.2	976	US-09-104-324B-4	Sequence 4, Appl
5	117	11.2	976	US-09-538-092-1339	Sequence 1339, Ap
6	114	10.9	284	US-09-914-259-55	Sequence 55, Appl
7	113.5	10.8	284	US-09-914-259-62	Sequence 62, Appl
8	112.5	10.7	630	US-09-248-796A-20275	Sequence 20275, A
9	112	10.7	284	US-09-914-259-43	Sequence 43, Appl
10	111	10.6	1937	US-09-538-092-918	Sequence 918, Appl
11	110	10.5	284	US-09-914-259-46	Sequence 46, Appl
12	110	10.5	284	US-09-914-259-49	Sequence 49, Appl
13	109.5	10.5	170	US-09-438-185A-708	Sequence 708, Appl
14	109	10.4	284	US-09-914-259-51	Sequence 51, Appl
15	108	10.3	817	US-09-248-796A-20276	Sequence 20276, A
16	107	10.2	281	US-09-914-259-63	Sequence 63, Appl
17	107	10.2	284	US-09-167-206-10	Sequence 10, Appl
18	107	10.2	284	US-09-914-259-40	Sequence 40, Appl
19	107	10.2	284	US-09-914-259-48	Sequence 48, Appl
20	107	10.2	284	US-09-914-259-50	Sequence 50, Appl
21	106.5	10.2	284	US-09-914-259-41	Sequence 41, Appl
22	106	10.1	534	US-09-103-664A-2	Sequence 2, Appl
23	106	10.1	1786	US-08-973-462-8	Sequence 16, Appl
24	105.5	10.1	224	US-08-272-255-16	Sequence 16, Appl
25	105.5	10.1	224	PCT-US95-08555-16	Sequence 16, Appl
26	105.5	10.1	245	US-10-164-595-34	Sequence 34, Appl
27	105	10.0	245	US-09-914-259-65	Sequence 65, Appl

28	105	10.0	251	4	US-09-914-259-64	Sequence 64, Appl
29	105	10.0	284	4	US-09-914-259-60	Sequence 60, Appl
30	104	9.9	1939	4	US-09-538-092-915	Sequence 915, Appl
31	104	9.9	1939	4	US-09-949-016-11104	Sequence 11104, A
32	102	9.7	281	4	US-09-914-259-45	Sequence 45, Appl
33	102	9.7	284	4	US-09-914-259-47	Sequence 47, Appl
34	102	9.7	284	4	US-09-914-259-57	Sequence 57, Appl
35	102	9.7	372	1	US-07-813-584A-3	Sequence 3, Appl
36	102	9.7	372	1	US-08-330-515-3	Sequence 3, Appl
37	102	9.7	1031	4	US-09-914-259-24	Sequence 24, Appl
38	102	9.7	1940	4	US-09-538-092-901	Sequence 901, Appl
39	102	9.7	1963	4	US-09-949-016-8888	Sequence 8888, Ap
40	101.5	9.7	2662	4	US-09-595-684B-31	Sequence 31, Appl
41	101.5	9.7	2663	4	US-09-538-092-1252	Sequence 1252, Ap
42	100.5	9.6	284	4	US-09-914-259-59	Sequence 59, Appl
43	100	9.6	1104	3	US-08-923-992A-4	Sequence 4, Appl
44	100	9.6	1164	3	US-08-923-992A-10	Sequence 10, Appl
45	99.5	9.5	588	4	US-08-714-741-42	Sequence 42, Appl

## ALIGNMENTS

```

RESULT 1
US-09-621-976-4959
; Sequence 4959, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Joubert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4959
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-4959

Query Match      56.7%; Score 593.5; DB 4; Length 127;
Best Local Similarity 89.6%; Pred. 1.7e-50;
Matches 120; Conservative 5; Mismatches 2; Indels 7; Gaps 1;

QY      1 MSKKGSLAEKRTMMEIFSETKVFOLEKIAPEKGTITAMSVKEVLQSLVDDGV 60
      |||
DB      1 MSKKGSLAEKRTMMEIFSETKVFOLEKIAPEKGTITAMSVKEVLQSLVDDGV 60

QY      61 DERIGTSNYVAFFSKALHARKHLEVLSEGLSGSQHSLQSIKAKIGRCETERR 120
      |||
DB      61 DERIGTSNYVAFFSKALHARKHLEVLSEGLSGSQHSLQSIKAKIGRCETERR 117

QY      121 TRLAKELSLRDPOR 134
      |||
DB      118 ----IKLSGMQBER 127

RESULT 2
US-09-248-796A-14485
; Sequence 14485, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13

```

PRIOR APPLICATION NUMBER: US 60/096,409  
 PRIOR FILING DATE: 1998-08-13  
 NUMBER OF SEQ ID NOS: 28208  
 SEQ ID NO: 14485  
 LENGTH: 260  
 TYPE: PRT  
 ORGANISM: Candida albicans  
 US-09-248-796A-14485

Query Match  
 Best Local Similarity 18.3%; Score 191.5; DB 4; Length 260;  
 Matches 59; Conservative 47; Mismatches 75; Indels 37; Gaps 8;

QY 1 MSKKGLSABEKRTMEIFSETKDVPOLKLEKIAPEKGTATMSVKEVLOSLVDGMDV 60  
 DB 60 MPKKGTLQEEKLSALWFOSDMEFTYLAKEIESKASKOCKIPPMOMELVALVEEGLV 60  
 QY 61 DCRIGTSNYWAPFSAKALHARKKLEVLSSQLSESGOKHABLOKSIKAK-----IGR 114  
 DB 120 EDDRGGTNNLWSPF-----YLHKKQ-----QETHRLNRTIANLETERDSLICH 165  
 QY 115 C-ETERRTRAKELSLR-----DOROLKAEVETKCDPOVEEIRQANKVAKKA 166  
 DB 166 CKDGTGVNQTHERAKIRFCQSLERIDSIQSLOSLKSES--VE-----NLVTSIAF 218  
 QY 167 NRTWDNIFAIKSWAKKFGFEENKIDRTFGIPEDFDYI 204  
 DB 219 --PSDSIDILICYLSHQGLTMTTKTEFFELPLEFEERI 254

RESULT 3

US-09-538-092-1321  
 Sequence 1321, Application US/09538092  
 Patent No. 6753314  
 GENERAL INFORMATION:  
 APPLICANT: Genetec, Inc.  
 APPLICANT: Manifest, Inc.  
 TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
 FILE REFERENCE: 15966-542  
 CURRENT APPLICATION NUMBER: US/09/538,092  
 PRIOR FILING DATE: 2000-03-29  
 PRIOR APPLICATION NUMBER: 60/127,352  
 PRIOR FILING DATE: 1999-04-01  
 PRIOR APPLICATION NUMBER: 60/178,965  
 NUMBER OF SEQ ID NOS: 1387  
 SOFTWARE: CurapatSeqFormat Version 0.9  
 SEQ ID NO: 1321  
 LENGTH: 1270  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (0)...(0)  
 OTHER INFORMATION: Polypeptide Accession Number, Q14203  
 US-09-538-092-1321

Query Match  
 Best Local Similarity 12.8%; Score 134.5; DB 4; Length 1270;  
 Matches 50; Conservative 43; Mismatches 106; Indels 15; Gaps 5;  
 QY 2 SKKGLSABEKRTMEIFSETKDVPOLKLEKIAPEKGTATMSVKEVLOSLVDGMDV 61  
 DB 271 ARKKEKALFAKERYMEADITADALEMTLDKEMAEERASLSQEVVALKERYDELTTD 330  
 QY 62 CERI-----GTSNYWAPFSAKALHARKKLEVLSSQLSESGOKHABLOKSIKAK 110  
 DB 331 LEILKALIEKSGDGAASSYOLKQLEEQVARKKALVNRDLSSEKOEHVKLQKLMER- 389  
 QY 111 KIGCE--ERTRAKELSLRDOROLKAEVETKCDPOVEEIRQANKVAKKA 166  
 DB 390 KNELEVQOQERLOEELSQAEISTIDELKQVDAALGAE-ENVENTLRNINLSEKVE 448

QY 169 WTDNIFAIKSWAKKFGFEENKIDRTFGIPEDFDYI 202  
 DB 449 LRETVGDLEANNENDELQENARFETELAEQLD 482

RESULT 4

US-09-104-324B-4  
 Sequence 4, Application US/09104324B  
 Patent No. 6332460  
 GENERAL INFORMATION:  
 APPLICANT: T rec1, Ozlem Sahin, Ugur, Pfreundschuh, Michael  
 TITLE OF INVENTION: Methods For Diagnosis And Treating Cancers,  
 TITLE OF INVENTION: And Methods For Identifying Pathogenic Markers In A Sample Of  
 NUMBER OF INVENTION: No. 6332460mal Cells  
 CORRESPONDENCE ADDRESSES: 4  
 ADDRESSER: Fulbright & Jaworski LLP  
 STREET: 666 Fifth Avenue  
 CITY: New York City  
 STATE: New York  
 ZIP: 10103  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
 COMPUTER: IBM  
 OPERATING SYSTEM: PC-DOS  
 SOFTWARE: Wordperfect  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/104,324B  
 FILING DATE: 25-June-1998  
 CLASSIFICATION: 435  
 PRIOR APPLICATION: 435  
 APPLICATION NUMBER: 08/892,702  
 FILING DATE: 15-July-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Hanson, No. 6332460man D.  
 REGISTRATION NUMBER: 30,946  
 TELEPHONE/DOCKET NUMBER: LUD 5491  
 TELECOMMUNICATION INFORMATION:  
 TELEFAX: (212) 752-5958  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 976 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 US-09-104-324B-4

Query Match  
 Best Local Similarity 11.2%; Score 117; DB 3; Length 976;  
 Matches 63; Conservative 38; Mismatches 85; Indels 78; Gaps 13;

QY 3 KKKGLSABE--KRTMEI--FSETKDY-----FQDKLEKIAPEKGTATMSVKEV 51  
 DB 413 QKSSSELEEMTKLNNKEVELEIKVLEKETTLYENKQPEKIAELKQ-TEQELIGLL 471  
 QY 52 QSL---VDDAMVDCERIGTSNYWAPFSAKALHARKKLEVLSSQLSESGOKHABLOKSIKAK 110  
 DB 472 QAREKEVHLELTQRTAITTSEQYKVEYKDLKTELNEKAKTELTSNCKLSTENKEIT 531  
 QY 91-----SOLSESGOKHABLOKSIKAKIGRGTETERRTRAKELSLRDOROLK 136  
 DB 532 QETSDMTLELKNQOEDINNKKOEERMLKQIE--NLQETETOLRNELEFVVEELKQKORDE 589  
 QY 137 LKAEVEKTKD-CD-----POVEEIRQANKVAKK--ANRTWDNIFAIK 166  
 DB 590 VCKRDKSSEENCNLRRQVENKRYKYLEEQOENKALKKGTABSKQALNVYIKVNLKLE 649  
 QY 178 -SWAKKFG-----PENKI 191  
 DB 650 LESAKQKFGELITTYOKETEDKDI 673

RESULT 5



US-09-538-092-1339  
 ; Sequence 1339, Application US/09538092  
 ; Patent No. 6753314  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Giot, Loic  
 ; APPLICANT: Mansfield, Traci A.  
 ; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
 ; FILE REFERENCE: 15966-542  
 ; CURRENT FILING DATE: 2000-03-29  
 ; PRIOR FILING DATE: 1999-04-01  
 ; PRIOR APPLICATION NUMBER: 60/127,352  
 ; PRIOR FILING DATE: 2000-02-01  
 ; NUMBER OF SEQ ID NOS: 1387  
 ; SOFTWARE: CuraPatSeqFormatter Version 0.9  
 ; SEQ ID NO 1339  
 ; LENGTH: 976  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (0)...(0)  
 ; OTHER INFORMATION: Polypeptide Accession Number Q15431  
 US-09-538-092-1339

Query Match 11.2%; Score 117; DB 4; Length 976;  
 Best Local Similarity 23.9%; Pred. No. 0.0099;  
 Matches 63; Conservative 38; Mismatches 85; Indels 78; Gaps 13;

QY 3 KKKGISAE--KRTMMELFSEITDV-----FOLKLEKIAPEKIGITAMSVKEVL 51  
 DB 413 OKKSELEBMTLTNNKEVELELKKVIGEKETLLYENKQFPIKIBELKG-TEQELIGL 471  
 QY 52 OSL--VDDGAWDCRIGTSNYWAFPSKAL-----HARKKLEVL- 90  
 DB 472 QAREKEVHDLQLTAITTSQYYSKEVDLKTLEENKLNTELTSHCNKLSLENKELT 531  
 QY 91 -----SOLSEGSOKHASLOKSIKAKIGRCETERTRLAKELSLDROEQ 136  
 DB 532 QETSMTELEKQOQDINNKKQERMLKQIE-NLOETITOLRLELVREELKQKDE 589  
 QY 137 LKAEEKYKD-CD-----POVVEIRQANKVAKA--AANRWTDNIFAIR----- 177  
 DB 590 VKCKLDKSEBNCNNLRKQVENNKYIEELQGENKALKKKGTASQKLANVEIKVKNLELE 649  
 QY 178 -SWAKRKG-----PEENKI 191  
 DB 650 LESAKQKFGIEITDYQKEIEDK 673

RESULT 6  
 US-09-914-259-55  
 ; Sequence 55, Application US/09914259  
 ; Patent No. 6495336  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Makowski, Lee  
 ; APPLICANT: Hyman, Paul  
 ; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES  
 ; FILE REFERENCE: 8471-010-999  
 ; CURRENT FILING DATE: 2000-11-21  
 ; NUMBER OF SEQ ID NOS: 180  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 55  
 ; LENGTH: 284  
 ; TYPE: PRT  
 ; ORGANISM: Sus scrofa  
 US-09-914-259-55

Query Match 10.9%; Score 114; DB 4; Length 284;  
 Best Local Similarity 23.2%; Pred. No. 0.0037;

Matches 52; Conservative 37; Mismatches 61; Indels 74; Gaps 11;  
 QY 1 MSKKGLSAE--EKRTMMELFSEITDVFOLEKLEKIAPEKIGITAMSVKEVLQSLVD 57  
 DB 73 LAEKATADAEADVASINRIQLFEEELDRAQ-----ERLA-----TALQKEAEKRADE 122  
 QY 58 ---GNVDCERIGTSNYWAFPSKALHAR---KHLVLESQSLSEGSQK----- 100  
 DB 123 SERGM-----KVIESRAQKDEERMEIOETOLKEA--KHIAEDDRKYE 163  
 QY 101 -----ASLOKSIKAKI--GRCTEERTRLAKELSLRDQREQLKAEVRYKQD 148  
 DB 164 EVARLVIIIESDLERAEERAEISEGKC-----AEIEBELKVTNNKLSLEAQAEKYSQKE 218  
 QY 149 POVVEIRQANKVAKAANRWTDNIFAISWAKRKGFEENKID 192  
 DB 219 DKYEIEIKVLSDKLKEAETR-----AEFRERSVTLEKSID 254

RESULT 7  
 US-09-914-259-62  
 ; Sequence 62, Application US/09914259  
 ; Patent No. 6495336  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Makowski, Lee  
 ; APPLICANT: Hyman, Paul  
 ; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES  
 ; FILE REFERENCE: 8471-010-999  
 ; CURRENT FILING DATE: 2000-11-21  
 ; NUMBER OF SEQ ID NOS: 180  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 62  
 ; LENGTH: 284  
 ; TYPE: PRT  
 ; ORGANISM: Rattus norvegicus  
 US-09-914-259-62

Query Match 10.8%; Score 113.5; DB 4; Length 284;  
 Best Local Similarity 21.1%; Pred. No. 0.0042;  
 Matches 48; Conservative 48; Mismatches 78; Indels 53; Gaps 8;

QY 2 SKKKGLSAE--EKRTMMELFSEITDVFOLEKLEKIAPEKIGITAMSVK-EVLQSLVDGM 59  
 DB 45 AKETLLRASDBDRRVLELHKAEDSLAAD-ETAKAEADVASINRIQLVEEELDRAQ 103  
 QY 60 VDCERIGTSNYWAFPSKALHARKKLEVLSEGSQK----- 99  
 DB 104 ---ERLATALQKLEAEKRADESERGMKVIESRAQKDEKMEIOETOLKEAHHIAEDADR 160  
 QY 100 -----HSLQSIKAKI--GRCTEERTRLAKELSLRDQREQLKAEVRYK 145  
 DB 161 KYEVARLVIIIESDLERAEERAEISEGKC-----AEIEBELKVTNNKLSLEAQAEKYS 215  
 QY 146 DCDPOVVEIRQANKVAKAANRWTDNIFAISWAKRKGFEENKID 192  
 DB 216 QKEDKYEIEIKVLSDKLKEAETR-----AEFRERSVTLEKSID 254

RESULT 8  
 US-09-248-796A-20275  
 ; Sequence 20275, Application US/09248796A  
 ; Patent No. 6747137  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Keith Weinstein et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 ; FILE REFERENCE: 107196.132  
 ; CURRENT FILING DATE: 1999-02-12  
 ; PRIOR APPLICATION NUMBER: US 60/074,725  
 ; PRIOR FILING DATE: 1998-02-13

Thu Mar 24 08:29:11 2005

us-10-087-190-3.ral

Page 4

PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 20275  
LENGTH: 630  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-20275

Query Match  
Best Local Similarity 10.7%; Score 112.5; DB 4; Length 630;  
Matches 49; Conservative 43; Mismatches 79; Indels 35; Gaps 8;

QY 3 KKKLSAEKRTMMEIFSETDVQOLDEKIAPEKGGIT-AMSKYVLSQVDDGV 60  
DB 158 KTKNSDTELEKQLELEKVA-----LDLQTADEKLGITTEKLEKLETVNSG-- 210  
QY 61 DCRIGTSNYWAFPS--KALHARKHLELV-----BSQSESGOKHSLQSKI----- 107  
DB 211 -LSTTSELALTKTVKSLSEKEKELOFLSKSKLELDYIOGHSDISEKALYDEL 266  
QY 108 -EKATIGCEETEERTLAKELSLDQREQLAEVEKYKCDPQVVEIRQANKVAKA 166  
DB 267 KEKTKQFDSDCKKLTLENDLSTKGLTEKTEKTSKFKNLEKKEIKVAKLELELK 326  
QY 167 NRTWDNIFAIKSWAKRFGFEENKID 192  
DB 327 N--DN-----SGAKTELEKVKSLK 344

RESULT 9

US-09-914-259-43  
Sequence 43; Application US/09914259  
Patent No. 6495336  
GENERAL INFORMATION:  
APPLICANT: Makowski, Lee  
APPLICANT: Hyman, Paul  
TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES  
FILE REFERENCE: 8471-010-999  
CURRENT FILING DATE: 2000-11-21  
NUMBER OF SEQ ID NOS: 180  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 43  
LENGTH: 284  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-914-259-43

Query Match  
Best Local Similarity 10.7%; Score 112; DB 4; Length 284;  
Matches 52; Conservative 41; Mismatches 72; Indels 56; Gaps 9;

QY 1 MSKKKGLSAEEK--RTMMEIFSETKDVFO-----LKOLEKIAPEKGGITAMSV 47  
DB 73 LAEKKAADAEVAVASLNRIQLVEBELDRAOERLATLQKLEAEKAADESERGMKVYEN 132  
QY 48 KEVLSQVDDGVDDCRIGTSNYWAFPSKALHARKHLELVESQSESGOKHSLQSKI 107  
DB 133 R-----ALNDEEKMELOEI-----OLKAGHIAEALRYEVAKLVIIIGDL 176  
QY 108 EKAKTGRCETEERTLA-----KELSLDQREQLAEVEKYKCDPQVVEIRQANK 159  
DB 177 EK-----TERAELAESKSELSEELNKNVNNIKSLKAEKISQKEDYEIEIKILT 229  
QY 160 KVAAEAAANRTWDNIFAIKSWAKRFGFEENKIDRTFGPIPED 200  
DB 230 DKLKEATRAE--FAERSVA-----KLEKTIIDLED 258

RESULT 10  
US-09-538-092-918

Sequence 918; Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:  
APPLICANT: Glot, Loic  
APPLICANT: Mansfield, Tracy A.  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US/09/538,092  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
NUMBER OF SEQ ID NOS: 60/178,965  
SOFTWARE: CuraPatSeqFormatter Version 0.9  
SEQ ID NO 918  
LENGTH: 1937  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (0).....(0)  
OTHER INFORMATION: Polyptide Accession Number P13535  
US-09-538-092-918

Query Match  
Best Local Similarity 10.6%; Score 111; DB 4; Length 1937;  
Matches 54; Conservative 42; Mismatches 88; Indels 50; Gaps 9;

QY 4 KKKLSAEKRTMMEIFSETDVQOLDEKIAPEKGGIT-AMSKYVLSQVDDGV 60  
DB 1361 ORALSKANSEVAVQKRTYETDAIORTLEAEKGLAQLEAEHVAVAKASIEKT 1420  
QY 48 KEVLSQVDDGVDDCRIGTS-----NYWAFPSKALHARKHLELVESQSESGO 98  
DB 1421 KORLQNEVEDLMDIVERSNAACAALDKQRNF-----DKVLEWKQKYEETOMAEASOK 1475  
QY 99 KHALQSIKAK-----IGRCET--ERTLAKELSLDQ-----RQQLAE 140  
DB 1476 ESRSLSTELFKVKNVVEESLDQLETLRENNKLOEFTSLDTEIAGGQIHELKIKQ 1535  
QY 141 VEKYKCDPQVVEIRQANKVAKA--ANRTWDNIFAIKSWAKRFGFEENKID 193  
DB 1536 VEOEK-CEIQALAEAEASLEHEGKITRIQLLELNQVSEVDKIAEDRIDQ 1588

RESULT 11

US-09-914-259-46  
Sequence 46; Application US/09914259  
Patent No. 6495336  
GENERAL INFORMATION:  
APPLICANT: Makowski, Lee  
APPLICANT: Hyman, Paul  
TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES  
FILE REFERENCE: 8471-010-999  
CURRENT FILING DATE: 2000-11-21  
NUMBER OF SEQ ID NOS: 180  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 46  
LENGTH: 284  
TYPE: PRT  
ORGANISM: Brachydanio rerio  
US-09-914-259-46

Query Match  
Best Local Similarity 10.5%; Score 110; DB 4; Length 284;  
Matches 51; Conservative 44; Mismatches 72; Indels 52; Gaps 10;

QY 1 MSKKKGLSAE--ERTMMEIFSETKDVFO-----LKOLEKIAPEKGGITAMSV 47  
DB 73 LAEKKAADAEVAVASLNRIQLVEBELDRAOERLATLQKLEAEKAADESERGMKVYEN 132

```

Query Match Similarity 10.5%; Score 110; DB 4; Length 284;
Best Local Similarity 23.2%; Pred. No. 0.0092;
Matches 48; Conservative 39; Mismatches 80; Indels 40; Gaps 8

QY      1 MSKKKGLSAE--EKRTMMEIFSETKDVQ-----LMDLEKIAPK-EKGITAMSV 47
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB      73 LAEKATDAEADVASINRRIGQVEEBELDRAQRLATALQLEAEKADSESGMKVIEN 132
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY      48 KEVLSLVDDGAVDCEERIGTSNYTAAFPKALHARGKLEVLSEQSBSGQKIASLQKSI 107
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB      133 R-----ALMDEEKIELOEI-----QKKAHGIAEADRKYEYARKVILIEGDI 176
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY      108 EKATIGCQETERR--TRLAKEISLRDQREOLKAIVEYKDCDPQVVEEIRQANKYAKKA 165
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB      177 ERAE-ERAEISSKCAELEEBELKTYNNIKSLDEAQEKYSQKEDKYEBELKVLTDYKKA 235
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY      166 ANRWTDNIFAIKSWAKRKFGFEENKID 192
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB      236 ETR-----AEFAERTYAKLEKSID 254
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

RESULT 13
US-09-438-185A-708
; Sequence 708, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08

```

```

RESULT 14
US-09-914-259-51
; Sequence 51. Application US/09914259
; Patent No. 649536
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-914-259-51

Query Match          10.4%; Score 109; DB 4; Length 284;
Best Local Similarity 23.7%; Pred. No. 0.011;
Matches 49; Conservative 37; Mismatches 81; Indels 40; Gaps 8;

QY      1 MSKKKGLSAE---EKTRMMEIFSETKDVPQ-----LKDLEKIAPK-EKGITAMSV 47
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      73 LSDKATDABSDVASLNRRIGQVLEEIDPRAOERISTLQKLEAEKKAADSEBGMKVIEN 132
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      48 KEVLQSLVDGMVDCERIGTSNYVWAPPSPKALHARKHKLTVLESQLESGSQHASLQKSI 107
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
          133 R---ALKDEKMKMLQEI-----QLKKAQIIAEADRAKREBVARKLVIIEGDL 176
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      108 EKAKIGCEFTBER--TLAAELSLRPQREQLKAEVEKYKDCDQVVEIIPQANKVAKEA 165
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
          177 ERAE-EAEAEISESKCALEBELKLTVMNLKSLAEQAKEYSQKEDKYEIEIKVLITDKLEA 235
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      166 ANRWTNIPAIKSWAKRKFGEENKID 192
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
          236 ETR-----AEFAERTVAKLEKSID 254
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
US-09-248-796A-20276
; Sequence 20276. Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
```

```

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; LENGTH: 817
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20276

```

```

Query Match      10.3% Score 108; DB 4; Length 817;
Best Local Similarity 24.4%; Pred. No. 0.059;
Matches 49; Conservative 43; Mismatches 85; Indels 24; Gaps 8;

```

```

QY 9 AAEKRTMMELFSETKDVQFQKDLKELKAPKEKITAMSVK--EVLQSLVD---DGMVDCE 63
DB 426 AAEKQV-LDEKKNODRIDTEAEKIAARKQELKLELAQAEKDEILKPTLDELKESAKLE 484
QY 64 RIGTSNYWAFPSKALH---ARKKLEVLSEQLSEGSOKHASLQKSIKAKIGRCETEE 119
DB 485 EVTNARDELANEYKASEDLNKEYEKEKLAELSKLOEAKNDIEKYTTDIEEA-----TAK 538
QY 120 RTRIAKELSLRDQEQKAVEK-YKDCDPQVVEEIRQANKKAKAKANRWTDNIFAIKS 178
DB 539 HESTDKVEAELOELHDKERKADAEKEHEDLDGKLELEKOKQJLHEDKATKKKKDILAAIDE 598
QY 179 WAKRKFGFEENKIDRTGFIPE 199
DB 539 KVK-----DEHKINSE--LPE 612

```

Search completed: March 23, 2005, 19:19:27  
Job time : 44 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model.

Run on: March 23, 2005, 19:17:58 ; Search time 138 seconds  
(without alignments)  
491.853 Million cell updates/sec

Title: US-10-087-190-3

Perfect score: 1047  
Sequence: 1 MSKKKGSLAEKRTTMEIF.....FEENKIDRTFGIPEDFDYID 205

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1407402 seqs, 311100923 residues

Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_NEW\_PUB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1047	100.0	205	US-09-799-250-2	Sequence 2, Appli
2	1047	100.0	205	US-10-087-190-3	Sequence 3, Appli
3	1047	100.0	205	US-10-087-190-14	Sequence 14, Appli
4	1047	100.0	205	US-10-087-190-20	Sequence 20, Appli
5	1047	100.0	205	US-10-087-190-21	Sequence 21, Appli
6	1047	100.0	205	US-10-087-190-22	Sequence 22, Appli
7	1047	100.0	205	US-10-087-190-61	Sequence 61, Appli
8	1047	100.0	205	US-10-408-765A-1821	Sequence 1821, Ap
9	1036.5	99.0	206	US-10-087-190-44	Sequence 44, Appli
10	1036.5	99.0	206	US-10-087-190-45	Sequence 45, Appli
11	1011	96.6	198	US-10-087-190-24	Sequence 24, Appli
12	975	93.1	190	US-10-087-190-13	Sequence 13, Appli
13	975	93.1	190	US-10-087-190-19	Sequence 19, Appli

14	975	93.1	190	US-10-087-190-67	Sequence 67, Appli
15	975	93.1	190	US-10-087-190-68	Sequence 68, Appli
16	975	93.1	190	US-10-087-190-69	Sequence 69, Appli
17	948	90.5	205	US-09-799-250-4	Sequence 4, Appli
18	948	90.5	205	US-10-087-190-23	Sequence 23, Appli
19	945.5	90.3	190	US-10-087-190-11	Sequence 11, Appli
20	945.5	90.3	190	US-10-087-190-18	Sequence 18, Appli
21	945.5	90.3	190	US-10-087-190-62	Sequence 62, Appli
22	945.5	90.3	190	US-10-087-190-53	Sequence 53, Appli
23	614	58.6	122	US-10-087-190-53	Sequence 9, Appli
24	591	56.4	122	US-10-087-190-9	Sequence 17, Appli
25	591	56.4	122	US-10-087-190-17	Sequence 54, Appli
26	591	56.4	122	US-10-087-190-54	Sequence 55, Appli
27	591	56.4	122	US-10-087-190-55	Sequence 7, Appli
28	584	55.8	119	US-10-087-190-7	Sequence 16, Appli
29	584	55.8	119	US-10-087-190-16	Sequence 47, Appli
30	584	55.8	119	US-10-087-190-47	Sequence 5, Appli
31	465	44.4	126	US-10-087-190-5	Sequence 15, Appli
32	465	44.4	126	US-10-087-190-15	Sequence 46, Appli
33	465	44.4	126	US-10-087-190-46	Sequence 162899,
34	427	40.8	207	US-10-437-963-162899	Sequence 162901,
35	394	37.6	235	US-09-799-250-5	Sequence 5, Appli
36	349	33.3	200	US-10-087-190-25	Sequence 25, Appli
37	349	33.3	200	US-10-087-190-35	Sequence 5443, Ap
38	287	27.4	79	US-10-106-698-5443	Sequence 13046,
39	255.5	24.4	218	US-10-083-357-1304	Sequence 2096, Ap
40	214.5	20.5	74	US-10-424-599-192650	Sequence 7646, Ap
41	121.5	11.6	715	US-10-408-765A-2096	Sequence 48, Appli
42	121.5	11.6	1881	US-10-032-585-7646	Sequence 1205, Ap
43	120.5	11.5	1798	US-09-981-151A-48	Sequence 34, Appli
44	117	11.2	795	US-10-408-765A-1205	
45	117	11.2	795	US-10-416-330-34	

## ALIGNMENTS

### RESULT 1

US-09-799-250-2  
; Sequence 2, Application US/09799250  
; Publication No. US20030032087A1  
; GENERAL INFORMATION:  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Arthur B. Raifano  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 121P1: A TISSUE SPECIFIC PROTEIN  
; TITLE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS  
; FILE REFERENCE: 129.34US01  
; CURRENT APPLICATION NUMBER: US/09/799,250  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 719  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 205  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-250-2

Query Match	100.0%	Score 1047;	DB 10;	Length 205;
Best Local Similarity	100.0%	Pred. No. 3e-78;		
Matches 205;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps	0;			
QY	1	MSKKKGSLAEKRTTMEIFSETKQVFLKDKETAPKEKGTITANSVKEVLSLVDDGAV	60	
DB	1	MSKKKGSLAEKRTTMEIFSETKQVFLKDKETAPKEKGTITANSVKEVLSLVDDGAV	60	
QY	61	DCERIGTSNYTAAFPSSKALHAKHKLTVLESQSLSGSQSHASLOKSIERAKIGRCETERR	120	
DB	61	DCERIGTSNYTAAFPSSKALHAKHKLTVLESQSLSGSQSHASLOKSIERAKIGRCETERR	120	

QY 121 TRAKELSSLRDQREQLKAEVEKYKDCDPQVVEEIRQANKVAKAANRWTDNIFAISMA 180  
 DB 121 TRAKELSSLRDQREQLKAEVEKYKDCDPQVVEEIRQANKVAKAANRWTDNIFAISMA 180  
 QY 181 KRKGFEEENKIDRTFGIPEDFDYID 205  
 DB 181 KRKGFEEENKIDRTFGIPEDFDYID 205

## RESULT 2

US-10-087-190-3  
 ; Sequence 3 Application US/10087190  
 ; Publication No. US2003022397A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Agensys, Inc.  
 ; APPLICANT: Chailita-Bid, Pia M.  
 ; APPLICANT: Hubert, Rene S.  
 ; APPLICANT: Raitano, Arthur B.  
 ; APPLICANT: Faris, Mary  
 ; APPLICANT: Afar, Daniel E. H.  
 ; APPLICANT: Ge, Wangmao  
 ; APPLICANT: Jakobovitz, Aya  
 ; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
 ; FILE REFERENCE: ENTITLED 121P1F1 USEFUL IN TREATMENT AND DETECTION OF CANCER  
 ; CURRENT APPLICATION NUMBER: US/10/087,190  
 ; PRIOR FILING DATE: 2003-01-28  
 ; PRIOR FILING DATE: 2001-03-05  
 ; NUMBER OF SEQ ID NOS: 69  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 205  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapiens  
 US-10-087-190-3

Query Match  
 Best Local Similarity 100.0%; Score 1047; DB 15; Length 205;  
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSAEERTRMWEIFSETDVQFQDKLEKIAPEKGTITANSVKEVLSLVDDGMV 60  
 DB 1 MSKKKGLSAEERTRMWEIFSETDVQFQDKLEKIAPEKGTITANSVKEVLSLVDDGMV 60  
 QY 61 DCRIGTSNYWAFPSKALHARKHKLVEVLSQSGSQKHSLOKSIKAKIGRCETEER 120  
 DB 61 DCRIGTSNYWAFPSKALHARKHKLVEVLSQSGSQKHSLOKSIKAKIGRCETEER 120  
 QY 121 TRAKELSSLRDQREQLKAEVEKYKDCDPQVVEEIRQANKVAKAANRWTDNIFAISMA 180  
 DB 121 TRAKELSSLRDQREQLKAEVEKYKDCDPQVVEEIRQANKVAKAANRWTDNIFAISMA 180  
 QY 181 KRKGFEEENKIDRTFGIPEDFDYID 205  
 DB 181 KRKGFEEENKIDRTFGIPEDFDYID 205

## RESULT 3

US-10-087-190-14  
 ; Sequence 14 Application US/10087190  
 ; Publication No. US2003022397A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Agensys, Inc.  
 ; APPLICANT: Chailita-Bid, Pia M.  
 ; APPLICANT: Hubert, Rene S.  
 ; APPLICANT: Raitano, Arthur B.  
 ; APPLICANT: Faris, Mary  
 ; APPLICANT: Afar, Daniel E. H.  
 ; APPLICANT: Ge, Wangmao  
 ; APPLICANT: Jakobovitz, Aya  
 ; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN

;; TITLE OF INVENTION: ENTITLED 121P1F1 USEFUL IN TREATMENT AND DETECTION OF CANCER  
 ;; FILE REFERENCE: 51158-20034.20  
 ;; CURRENT APPLICATION NUMBER: US/10/087,190  
 ;; PRIOR FILING DATE: 2003-01-28  
 ;; PRIOR FILING DATE: 2001-03-05  
 ;; NUMBER OF SEQ ID NOS: 69  
 ;; SOFTWARE: FastSeq for Windows Version 4.0  
 ;; SEQ ID NO 14  
 ;; LENGTH: 205  
 ;; TYPE: PRT  
 ;; ORGANISM: Homo Sapiens  
 US-10-087-190-14

Query Match  
 Best Local Similarity 100.0%; Score 1047; DB 15; Length 205;  
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSAEERTRMWEIFSETDVQFQDKLEKIAPEKGTITANSVKEVLSLVDDGMV 60  
 DB 1 MSKKKGLSAEERTRMWEIFSETDVQFQDKLEKIAPEKGTITANSVKEVLSLVDDGMV 60  
 QY 61 DCRIGTSNYWAFPSKALHARKHKLVEVLSQSGSQKHSLOKSIKAKIGRCETEER 120  
 DB 61 DCRIGTSNYWAFPSKALHARKHKLVEVLSQSGSQKHSLOKSIKAKIGRCETEER 120  
 QY 121 TRAKELSSLRDQREQLKAEVEKYKDCDPQVVEEIRQANKVAKAANRWTDNIFAISMA 180  
 DB 121 TRAKELSSLRDQREQLKAEVEKYKDCDPQVVEEIRQANKVAKAANRWTDNIFAISMA 180  
 QY 181 KRKGFEEENKIDRTFGIPEDFDYID 205  
 DB 181 KRKGFEEENKIDRTFGIPEDFDYID 205

## RESULT 4

US-10-087-190-20  
 ; Sequence 20 Application US/10087190  
 ; Publication No. US2003022397A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Agensys, Inc.  
 ; APPLICANT: Chailita-Bid, Pia M.  
 ; APPLICANT: Hubert, Rene S.  
 ; APPLICANT: Raitano, Arthur B.  
 ; APPLICANT: Faris, Mary  
 ; APPLICANT: Afar, Daniel E. H.  
 ; APPLICANT: Ge, Wangmao  
 ; APPLICANT: Jakobovitz, Aya  
 ; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
 ; FILE REFERENCE: ENTITLED 121P1F1 USEFUL IN TREATMENT AND DETECTION OF CANCER  
 ; CURRENT APPLICATION NUMBER: US/10/087,190  
 ; PRIOR FILING DATE: 2003-01-28  
 ; PRIOR FILING DATE: 2001-03-05  
 ; NUMBER OF SEQ ID NOS: 69  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 20  
 ; LENGTH: 205  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapiens  
 US-10-087-190-20

Query Match  
 Best Local Similarity 100.0%; Score 1047; DB 15; Length 205;  
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSAEERTRMWEIFSETDVQFQDKLEKIAPEKGTITANSVKEVLSLVDDGMV 60  
 DB 1 MSKKKGLSAEERTRMWEIFSETDVQFQDKLEKIAPEKGTITANSVKEVLSLVDDGMV 60  
 QY 61 DCRIGTSNYWAFPSKALHARKHKLVEVLSQSGSQKHSLOKSIKAKIGRCETEER 120  
 DB 61 DCRIGTSNYWAFPSKALHARKHKLVEVLSQSGSQKHSLOKSIKAKIGRCETEER 120



Db 61 DCRIGTSNYWAFPSKALHARKHLEVLSEQLSEGSQGHASLOKSIEKAKIGRCCTEER 120  
QY 121 TRLAEKLSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKKANRWTDNIFAIXSWA 180  
Db 121 TRLAEKLSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKKANRWTDNIFAIXSWA 180  
QY 181 KRKFGFEENKIDRTFGIPEDFDYID 205  
Db 181 KRKFGFEENKIDRTFGIPEDFDYID 205

## RESULT 5

US-10-087-190-21  
; Sequence 21, Application US/10087190  
; Publication No. US20030223997A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc.  
; APPLICANT: Challita-Bid, Pia M.  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Paris, Mary  
; APPLICANT: Afar, Daniel E. H.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovitz, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20034.20  
; CURRENT APPLICATION NUMBER: US/10/087,190  
; PRIOR FILING DATE: 2003-01-28  
; PRIOR APPLICATION NUMBER: US 09/779,250  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 205  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-087-190-21

Query Match 100.0%; Score 1047; DB 15; Length 205;  
Best Local Similarity 100.0%; Pred. No. 3e-78;  
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSKKKGLSABEKRTMEIFSETKDVFOUKDEKIAPEKKGITANSVKEVLQSLVDDGNV 60  
Db 1 MSKKKGLSABEKRTMEIFSETKDVFOUKDEKIAPEKKGITANSVKEVLQSLVDDGNV 60  
QY 61 DCRIGTSNYWAFPSKALHARKHLEVLSEQLSEGSQGHASLOKSIEKAKIGRCCTEER 120  
Db 61 DCRIGTSNYWAFPSKALHARKHLEVLSEQLSEGSQGHASLOKSIEKAKIGRCCTEER 120  
QY 121 TRLAEKLSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKKANRWTDNIFAIXSWA 180  
Db 121 TRLAEKLSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKKANRWTDNIFAIXSWA 180  
QY 181 KRKFGFEENKIDRTFGIPEDFDYID 205  
Db 181 KRKFGFEENKIDRTFGIPEDFDYID 205

## RESULT 6

US-10-087-190-22  
; Sequence 22, Application US/10087190  
; Publication No. US20030223997A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc.  
; APPLICANT: Challita-Bid, Pia M.  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Paris, Mary  
; APPLICANT: Afar, Daniel E. H.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovitz, Aya

; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20034.20  
; CURRENT APPLICATION NUMBER: US/10/087,190  
; PRIOR FILING DATE: 2003-01-28  
; PRIOR APPLICATION NUMBER: US 09/779,250  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 205  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-087-190-22

Query Match 100.0%; Score 1047; DB 15; Length 205;  
Best Local Similarity 100.0%; Pred. No. 3e-78;  
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSKKKGLSABEKRTMEIFSETKDVFOUKDEKIAPEKKGITANSVKEVLQSLVDDGNV 60  
Db 1 MSKKKGLSABEKRTMEIFSETKDVFOUKDEKIAPEKKGITANSVKEVLQSLVDDGNV 60  
QY 61 DCRIGTSNYWAFPSKALHARKHLEVLSEQLSEGSQGHASLOKSIEKAKIGRCCTEER 120  
Db 61 DCRIGTSNYWAFPSKALHARKHLEVLSEQLSEGSQGHASLOKSIEKAKIGRCCTEER 120  
QY 121 TRLAEKLSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKKANRWTDNIFAIXSWA 180  
Db 121 TRLAEKLSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKKANRWTDNIFAIXSWA 180  
QY 181 KRKFGFEENKIDRTFGIPEDFDYID 205  
Db 181 KRKFGFEENKIDRTFGIPEDFDYID 205

## RESULT 7

US-10-087-190-61  
; Sequence 61, Application US/10087190  
; Publication No. US20030223997A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc.  
; APPLICANT: Challita-Bid, Pia M.  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Paris, Mary  
; APPLICANT: Afar, Daniel E. H.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovitz, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20034.20  
; CURRENT APPLICATION NUMBER: US/10/087,190  
; PRIOR FILING DATE: 2003-01-28  
; PRIOR APPLICATION NUMBER: US 09/779,250  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 61  
; LENGTH: 205  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-087-190-61

Query Match 100.0%; Score 1047; DB 15; Length 205;  
Best Local Similarity 100.0%; Pred. No. 3e-78;  
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSKKKGLSABEKRTMEIFSETKDVFOUKDEKIAPEKKGITANSVKEVLQSLVDDGNV 60  
Db 1 MSKKKGLSABEKRTMEIFSETKDVFOUKDEKIAPEKKGITANSVKEVLQSLVDDGNV 60  
QY 61 DCRIGTSNYWAFPSKALHARKHLEVLSEQLSEGSQGHASLOKSIEKAKIGRCCTEER 120

```

Db      61 DCRIGTSNYWAFPSKALHARKHKLVELESQLSSEGSQKASLOKSIKAKIGRCCTEER 120
Qy      121 TRLAKELSLRDQREOLKAEVEKYKDCDPQVVEIRQANKVAKKANRWTDNIFAIKSWA 180
Db      121 TRLAKELSLRDQREOLKAEVEKYKDCDPQVVEIRQANKVAKKANRWTDNIFAIKSWA 180
Qy      181 KRKGFPEENKIDRTFGIPEDFDYID 205
Db      181 KRKGFPEENKIDRTFGIPEDFDYID 205

```

```

RESULT 8
US-10-408-765A-1821
; Sequence 1821, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Bo D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Martock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408.765A
; SOFTWARE: FASTSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 3077
; SEQ ID NO 1821
; LENGTH: 205
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-408-765A-1821

```

```

Query Match
Best Local Similarity 100.0%; Score 1047; DB 16; Length 205;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 MSKKKGLSAEKKRTMEIFSETKDVFOLEKIAKPKKGTANSVKEVLOSLVDGMV 60
Db      1 MSKKKGLSAEKKRTMEIFSETKDVFOLEKIAKPKKGTANSVKEVLOSLVDGMV 60
Qy      61 DCRIGTSNYWAFPSKALHARKHKLVELESQLSSEGSQKASLOKSIKAKIGRCCTEER 120
Db      61 DCRIGTSNYWAFPSKALHARKHKLVELESQLSSEGSQKASLOKSIKAKIGRCCTEER 120
Qy      121 TRLAKELSLRDQREOLKAEVEKYKDCDPQVVEIRQANKVAKKANRWTDNIFAIKSWA 180
Db      121 TRLAKELSLRDQREOLKAEVEKYKDCDPQVVEIRQANKVAKKANRWTDNIFAIKSWA 180
Qy      181 KRKGFPEENKIDRTFGIPEDFDYID 205
Db      181 KRKGFPEENKIDRTFGIPEDFDYID 205

```

```

RESULT 9
US-10-087-190-44
; Sequence 44, Application US/10087190
; Publication No. US2003022397A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Afari, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; ENTITLED 121P1 USEFUL IN TREATMENT AND DETECTION OF CANCER

```

```

; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087.190
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 206
; TYPE: PRF
; ORGANISM: Homo Sapiens
US-10-087-190-44

```

```

Query Match
Best Local Similarity 99.0%; Score 1036.5; DB 15; Length 206;
Matches 205; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy      1 MSKKKGLSAEKKRTMEIFSETKDVFOLEKIAKPKKGTANSVKEVLOSLVDGMV 60
Db      1 MSKKKGLSAEKKRTMEIFSETKDVFOLEKIAKPKKGTANSVKEVLOSLVDGMV 60
Qy      61 DCRIGTSNYWAFPSKALHARKHKLVELESQLSSEGSQKASLOKSIKAKIGRCCTEER 119
Db      61 DCRIGTSNYWAFPSKALHARKHKLVELESQLSSEGSQKASLOKSIKAKIGRCCTEER 119
Qy      120 RTRAKELSLRDQREOLKAEVEKYKDCDPQVVEIRQANKVAKKANRWTDNIFAIKSWA 179
Db      121 TRLAKELSLRDQREOLKAEVEKYKDCDPQVVEIRQANKVAKKANRWTDNIFAIKSWA 180
Qy      180 AKRKGFEENKIDRTFGIPEDFDYID 205
Db      181 AKRKGFEENKIDRTFGIPEDFDYID 206

```

```

RESULT 10
US-10-087-190-45
; Sequence 45, Application US/10087190
; Publication No. US2003022397A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Afari, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087.190
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 206
; TYPE: PRF
; ORGANISM: Homo Sapiens
US-10-087-190-45

```

```

Query Match
Best Local Similarity 99.0%; Score 1036.5; DB 15; Length 206;
Matches 205; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy      1 MSKKKGLSAEKKRTMEIFSETKDVFOLEKIAKPKKGTANSVKEVLOSLVDGMV 60
Db      1 MSKKKGLSAEKKRTMEIFSETKDVFOLEKIAKPKKGTANSVKEVLOSLVDGMV 60
Qy      61 DCRIGTSNYWAFPSKALHARKHKLVELESQLSSEGSQKASLOKSIKAKIGRCCTEER 119
Db      61 DCRIGTSNYWAFPSKALHARKHKLVELESQLSSEGSQKASLOKSIKAKIGRCCTEER 120

```



Db 61 SKALHARKHLEEVLESQSGOKHASLOKSIIEKAKIGRCETEERTRLAKELSLRDQRE 120  
QY 136 QKAEVEKYKDCDPQVEEIRQANKVAKAANRWTDNIFAIKSWAKRKGFEENKIDRTF 195  
Db 121 QKAEVEKYKDCDPQVEEIRQANKVAKAANRWTDNIFAIKSWAKRKGFEENKIDRTF 195  
QY 196 GIPEDFDYID 205  
Db 181 GIPEDFDYID 190

## RESULT 14

US-10-087-190-67  
; Sequence 67, Application US/10087190  
; Publication No. US2003023997A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc.  
; APPLICANT: Chailita-Bid, Pia M.  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Paris, Mary  
; APPLICANT: Afar, Daniel E. H.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovitz, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20034.20  
; CURRENT APPLICATION NUMBER: US/10/087,190  
; PRIOR FILING DATE: 2003-01-28  
; PRIOR APPLICATION NUMBER: US 09/779,250  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 67  
; LENGTH: 190  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-087-190-67

## Query Match

Best Local Similarity 93.1%; Score 975; DB 15; Length 190;  
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MMEIFSETKDVFOKDKLEKIAPEKKGITAMSVKEVLOSLVDGMVDCERIGTSNYMAFP 75  
Db 1 MMEIFSETKDVFOKDKLEKIAPEKKGITAMSVKEVLOSLVDGMVDCERIGTSNYMAFP 75  
QY 76 SKALHARKHLEEVLESQSGOKHASLOKSIIEKAKIGRCETEERTRLAKELSLRDQRE 135  
Db 61 SKALHARKHLEEVLESQSGOKHASLOKSIIEKAKIGRCETEERTRLAKELSLRDQRE 135  
QY 136 QKAEVEKYKDCDPQVEEIRQANKVAKAANRWTDNIFAIKSWAKRKGFEENKIDRTF 195  
Db 121 QKAEVEKYKDCDPQVEEIRQANKVAKAANRWTDNIFAIKSWAKRKGFEENKIDRTF 195  
QY 196 GIPEDFDYID 205  
Db 181 GIPEDFDYID 190

## RESULT 15

US-10-087-190-68  
; Sequence 68, Application US/10087190  
; Publication No. US2003023997A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc.  
; APPLICANT: Chailita-Bid, Pia M.  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Paris, Mary  
; APPLICANT: Afar, Daniel E. H.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovitz, Aya

; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20034.20  
; CURRENT APPLICATION NUMBER: US/10/087,190  
; PRIOR FILING DATE: 2003-01-28  
; PRIOR APPLICATION NUMBER: US 09/779,250  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 68  
; LENGTH: 190  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-087-190-68

## Query Match

Best Local Similarity 93.1%; Score 975; DB 15; Length 190;  
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MMEIFSETKDVFOKDKLEKIAPEKKGITAMSVKEVLOSLVDGMVDCERIGTSNYMAFP 75  
Db 1 MMEIFSETKDVFOKDKLEKIAPEKKGITAMSVKEVLOSLVDGMVDCERIGTSNYMAFP 75  
QY 76 SKALHARKHLEEVLESQSGOKHASLOKSIIEKAKIGRCETEERTRLAKELSLRDQRE 135  
Db 61 SKALHARKHLEEVLESQSGOKHASLOKSIIEKAKIGRCETEERTRLAKELSLRDQRE 135  
QY 136 QKAEVEKYKDCDPQVEEIRQANKVAKAANRWTDNIFAIKSWAKRKGFEENKIDRTF 195  
Db 121 QKAEVEKYKDCDPQVEEIRQANKVAKAANRWTDNIFAIKSWAKRKGFEENKIDRTF 195  
QY 196 GIPEDFDYID 205  
Db 181 GIPEDFDYID 190

Search completed: March 23, 2005, 19:30:49  
Job time : 139 secs



A/Residues: 1-128 <BEV>  
 A/Cross-references: UNIPROT:O9SZES, EMBL:AL078470, GSPDB:GN00062, ATSP:F19B15.200  
 A/Experimental source: cultivar Columbia, BAC clone F19B15  
 A/Genetic: A:Gene: A:SP:F19B15.200  
 A/Map position: 4  
 A/Introns: 13/3, 52/3, 88/3, 109/3

## Query Match

Best Local Similarity 18.8%; Score 197, DB 2, Length 128;  
 Matches 49, Conservative 23, Mismatches 37, Indels 14, Gaps 3,  
 Oy 16 MEISETVDVFOQLDKLRIKPKGTTAMSVKVLQSLVDGMDCEKIGTSNYWAF 75  
 Db 1 MLIETESQDFELKEKPKK-KGVISQSVADVQSLVDDVAADKIGIS----- 52  
 Oy 76 STALHARKKLEVESQSLSGQKHAQLQKSEKAKIRCTERTLAKELSLRDQRE 135  
 Db 53 ---LRSVQK---LESIDLQGSNKRRLAEVLVDQCEALMKKGRSEERTLALTKDLEKHK 106  
 Oy 136 QLK 138  
 Db 107 DLK 109

## RESULT 3

Se1134  
 N/Hypothetical protein YGL183c - Yeast (Saccharomyces cerevisiae)  
 C/Species: Saccharomyces cerevisiae  
 C/Date: 23-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 09-Jul-2004  
 A/Accession: Se1134; S64200  
 R/Bertani, I.; Cogillevina, M.; Zaccaria, P.; Kluma, R.; Bruchti, C.V.  
 Submitted to the EMBL Data Library, September 1995.  
 A/Description: The sequence analysis of a 7.9 kb DNA fragment from the left arm of S. cerevisiae new genes.  
 A/Reference number: S61128  
 A/Accession: Se1134  
 A/Molecule type: DNA  
 A/Residues: 1-174 <BER>  
 A/Cross-references: UNIPROT:P53102, EMBL:X91489, NID:G1143557, PIDN:CAA62791.1, PID:el995  
 R/Bruchti, C.V.; Cogillevina, M.; Bertani, I.; Kluma, R.; Zaccaria, P.; Delneri, D.  
 Submitted to the Protein Sequence Database, May 1996  
 A/Reference number: S64183  
 A/Accession: S64200  
 A/Molecule type: DNA  
 A/Residues: 1-174 <BRU>  
 A/Cross-references: EMBL:Z72705, NID:G1322796, PIDN:CAA6895.1, PID:el43495, PID:G132279  
 C/Experimental source: strain S288C  
 A/Genetic: A:Gene: SGD:MND1  
 A/Cross-references: SGD:S0003151  
 A/Map position: 7L

## Query Match

Best Local Similarity 16.1%; Score 168.5, DB 2, Length 174;  
 Matches 46, Conservative 42, Mismatches 68, Indels 13, Gaps 5,  
 Oy 45 MSVEVLQSLVD-DGAVDCEKIGTSNYWAFPSKALHARKKLEVESQSLSGQKHA 103  
 Db 1 MIVYDVLQVQWIDEDVIVSEKGNINITYWCPKQIOLQKYSDESEIKKIDEVKCDIATY 60  
 Oy 104 QKSIKRA-KIGCEETERTLAKELSLRDQREQLAEVEKXKDDPOVEIR----- 156  
 Db 61 KQELDKTALATGRKKFTVQGSYRREALLKPKKIDDEIKK-KNSLQKIESIRMAAKI 119  
 Oy 157 QANK-----VAKAANRWTDNIPAIKSMARKGFEENKIDRTFGIPDF 201  
 Db 120 QENQOIRLKKVHLKTKTDIILDIYKKFKPKQIRREIRGIEEP 168

## RESULT 4

JC5368

dynactin 1 - mouse

N/Alternate names: p150 Glued  
 C/Species: Mus musculus (house mouse)  
 C/Date: 28-May-1997 #sequence\_revision 18-Jul-1997 #text\_change 09-Jul-2004  
 A/Accession: JC5368  
 R/Jiang, W.; Weber, U.S.; Tokito, M.K.; Holzbaur, E.L.F.; Meister, M.H.  
 Biochem. Biophys. Res. Commun. 231, 344-347, 1997  
 A/Title: Mouse p150Glued (dynactin 1) cDNA sequence and evaluation as a candidate for  
 A/Reference number: JC5368, MUID:9723454, PMID:9070275  
 A/Accession: UC5368  
 A/Molecule type: mRNA  
 A/Residues: 1-1281 <JAN>  
 A/Cross-references: UNIPROT:O08788, GB:U60312, NID:G2104494, PIDN:AB57773.1, PID:G2  
 A/Experimental source: brain  
 C/Comment: This protein is a member of the oligomeric dynactin complex that is required for  
 A/Genetic: A:Gene: Dcn1  
 A/Map position: 6

## Query Match

Best Local Similarity 12.9%; Score 135.5, DB 2, Length 1281;  
 Matches 50, Conservative 43, Mismatches 106, Indels 15, Gaps 5,  
 Oy 2 SKKGLSAEKKTRMEISETKDVFOQLDKLRIKPKGTTAMSVKVLQSLVDGMD 61  
 Db 279 ARKAKALBAKRYMEWADYADIAIEMATLDEKMAEERBSIQDEVALKERVDLTTD 338  
 Oy 62 CERL-----GTSNYWAFPSKALHAR-KHKLEVESQSLSGQKHAQLQKSEKAKIR 110  
 Db 339 LETIKARIEKSDAASSYOLKQLEQONARLKDMLVRRLDSSSEKQEHVKQKMKER- 397  
 Oy 111 KIRCEET-ERTTIAKELSLRDQREQLKAEVEKXKDDPOVEIRQANKAKAANR 168  
 Db 398 KQELLEVQQRERLOEBSLOAESTIDELKEQVDALGAE-EMVEMLTDRNLLEKRYE 456  
 Oy 169 WTDNIPAIKSMARKGFEENKIDRTFGIPDFD 202  
 Db 457 LRETVGLLEAMNEMNDXLOENARRELELRDOLD 490

## RESULT 5

A41642  
 N/dynactin - chicken  
 C/Species: Gallus gallus (chicken)  
 C/Date: 28-Aug-1992 #sequence\_revision 28-Aug-1992 #text\_change 09-Jul-2004  
 A/Accession: A41642  
 R/Gill, S.R.; Schroer, T.A.; Szilak, I.; Steiner, E.R.; Sheetz, M.P.; Cleveland, D.W.  
 J. Cell Biol. 115, 1639-1650, 1991  
 A/Title: Dynactin, a conserved, ubiquitously expressed component of an activator of vesicular  
 A/Reference number: A41642, MUID:92098576, PMID:1836789  
 A/Status: Preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-1053 <GIL>  
 A/Cross-references: UNIPROT:P35458, GB:X62773, NID:G63920, PID:G63921  
 C/Keywords: cytoskeleton

## Query Match

Best Local Similarity 12.5%; Score 130.5, DB 2, Length 1053;  
 Matches 53, Conservative 42, Mismatches 93, Indels 37, Gaps 7,  
 Oy 2 SKKGLSAEKKTRMEISETKDVFOQLDKLRIKPKGTTAMSVKVLQSLVDGMD 57  
 Db 101 AKKAKALBAKRYMEWADYADIAIEMATLDEKMAEERBSIQDEVALKERVDLTTD 160  
 Oy 58 GMV-----DCERIGTSNYWAFPSKALHARKKLE-----VLSQSLSGQKHAQLQKSEKAKIR 110  
 Db 161 LETIKARIEKSDAASSYOLKQLEQONARLKDMLVRRLDSSSEKQEHVKQKMKER- 220  
 Oy 111 KIRCEETERTLAKELSLRDQREQLKAEVEKXKDDPOVEIRQANKAKAANR 168  
 Db 221 N-----TELESLRQQRERLOEBSLOAESTIDELKEQVDALGAEEMVEMLTDRNLLEKRYE 267



Qy 158 ANKVAEAAARNTDNIFAIKSWAKRRKFGFEENKIDRTFGIPEDFD 202  
 Db 268 RNLDEEKREURELFTVGDEANMENMDEIQENARETELELRQLD 312

## RESULT 6

kinectin 1 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #ext\_change 09-Jul-2004  
 C:Accession: S33763; 137947  
 R:Kruppa, G.; Puetterer, A.; Lemke, H.; Kronke, M.  
 submitted to the EMBL Data Library, April 1993  
 A:Description: Cloning and characterization of TAF, a novel transactivating protein.  
 A:Reference number: S33763  
 A:Accession: S33763  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1356 <KRU>  
 A:Cross-references: UNIPROT:Q86DUP2; EMBL:222551  
 R:Putterer, A.; Kruppa, G.; Kramer, B.; Lemke, H.; Kronke, M.  
 Mol. Biol. Cell 6, 161-170, 1995  
 A:Title: Molecular cloning and characterization of human kinectin.  
 A:Reference number: 137947; MUID:95306853; PMID:7787243  
 A:Accession: 137947  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1356 <RBS>  
 A:Cross-references: EMBL:222551; NID:g296163; PIDN:CAA80271.1; PID:g296164  
 C:Genetics:  
 A:Gene: GDB:KTN1  
 A:Cross-references: GDB:6165852; OMIM:600653

Query Match 11.6%; Score 121.5; DB 2; Length 1356;  
 Best Local Similarity 22.7%; Pred. No. 2.3;  
 Matches 50; Conservative 31; Mismatches 88; Indels 51; Gaps 5;

Qy 3 KKKGLSAEKKRTMEIFSETKDVLPOLKLEKIAPEKKGITA--MSKEVLQSLVDGAV 60  
 Db 1025 RKNNDLREKMEAEALASTEKMDKNTKSKKQOQVEALEAKVLLKLPKAVS 1084  
 Qy 61 DCEKIGTSNYWAPPSKALH---ARKKLEVLSEQLSEGSQKIASLQKSIKAKIGRC 115  
 Db 1085 P-SNLSYCGWHLGFEKKAKCMAAGTSGSEVKVLEHKLKADBMHTLLQJCEKYSVLA 1143  
 Qy 116 ETE-----BRTFLA-KELSSLD 132  
 Db 1144 ETEGILLQKLSRVEQENKVKVDESHKTIKQMOSSFTSSQELERLSKNKDIENLR 1203  
 Qy 133 QREQLAEVEKYKDCDPQVEEIRQANKVAEAAARNTDN 172  
 Db 1204 EREHLEMELEKAMERSTYTVRELRKDLTLTLOKLDLDS 1243

## RESULT 7

hypothetical protein At2g39300 [imported] - Arabidopsis thaliana  
 N:Alternate names: hypothetical protein T16B24.6  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #ext\_change 09-Jul-2004  
 C:Accession: T02572; F84815  
 R:Roundley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,  
 submitted to the EMBL Data Library, August 1998  
 A:Description: Arabidopsis thaliana chromosome II BAC T16B24 genomic sequence.  
 A:Reference number: Z14679  
 A:Accession: T02572  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-768 <ROU>  
 A:Cross-references: UNIPROT:Q08951; EMBL:AC004697; NID:g3402671; PID:g3402677  
 A:Experimental source: cultivar Columbia  
 R:Lin, X.; Kaul, S.; Roundley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

ues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:1061197  
 A:Accession: F84815  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-768 <STO>  
 C:Genetics:  
 A:Gene: T16B24.6; At2g39300  
 A:Map position: 2  
 A:Introns: 80/2; 665/3

Query Match 11.2%; Score 117; DB 2; Length 768;  
 Best Local Similarity 22.1%; Pred. No. 2.5;  
 Matches 50; Conservative 46; Mismatches 90; Indels 40; Gaps 9;

Qy 1 MSKKGLSANE-----EKRTMEIFSETK-----DVFOI-KOLEKIAPEKGIT 43  
 Db 247 MCKEDDVSELEKRYKEAKRYKVLSEEMEEKFLSDCPDIFSSLVGDIRQMEERVGL 305  
 Qy 44 MSYKVLQSLVDGMDCEKIGTSNYWAPPSKALHARKHLEV-LSQSLSEGSQKAS 102  
 Db 306 AFEVLSLRQWDERASTREDIRRVKNPMDLLKRLKEKTELQVLETLDRSSEWTS 365  
 Qy 103 LQKSEKAKIGRCETREERTLAKELSLRDQEQLEAVEKYKDCDPQVEEIRQANKVA 162  
 Db 366 ---KVESFKV-----EEKRLERVRLEAHNVSLQREISTHEKETERIDIRHLETV 416  
 Qy 163 KE---AANRWTDNIFAISWAKRRKFGFEENKIDRTFGIPEDFDYI 204  
 Db 417 TELSATAEEMRENFLLMQNLSKIQESYT-----GSTDDLVDYV 454

## RESULT 8

centrosome associated protein CEP250 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #ext\_change 21-Jul-2000  
 C:Accession: T08621  
 R:MacK, G.J.; Rees, J.; Sandilow, O.; Balczon, R.; Fritzler, M.J.; Rattner, J.B.  
 Arthritis Rheum. 41, 551-558, 1998  
 A:Title: Autoantibodies to a group of centrosomal proteins in human autoimmune sera reac  
 A:Reference number: Z16462; MUID:98165428; PMID:9506584  
 A:Accession: T08621  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-2442 <MAC>  
 A:Cross-references: EMBL:AF022655; NID:g2832236; PIDN:AAC06349.1; PID:g2832237  
 A:Experimental source: cell line HeLa

Query Match 11.1%; Score 116.5; DB 2; Length 2442;  
 Best Local Similarity 19.7%; Pred. No. 9.4;  
 Matches 46; Conservative 47; Mismatches 64; Indels 77; Gaps 8;

Qy 1 MSKKGLSAEKKRTMEIFSETKDVLPOLKDE--KIAPEKIGTAMSVKXVLSVLDG 58  
 Db 1662 LQKRIQVLEDDRTQTKLES-----DLQIKLSUREKRGELTTQROLOERAEBSG 1713  
 Qy 59 M-----VDCR-----RIGTSNYW 72  
 Db 1714 KQPSKAKQSGSLHMKLLIRDKKEVECCQENHHEQLKDLQEQLOGHRRVGET---- 1769  
 Qy 73 AFPSPALHARKHLEVLSEQLSE---GSQKIASLQKSIKAKIGRCETERTRLAKELS 128  
 Db 1770 ---SLLSQREDEIVLQOQLQEAEEQSELKQSLQSLQDEKRALARDQ-----ELE 1820  
 Qy 129 SLRDQEQLEAVEKYKDCDPQVEEIRQANKVAEAAARNTDNIFAISWAKR 182  
 Db 1821 ALQEQEQQAQGOEERKAKADALQALADQAHWTLLKERHGELODH---KEQARR 1870

## RESULT 9

F75103  
 conserved hypothetical protein PAB0812 - Pyrococcus abyssi (strain Orsay)  
 C/Species: Pyrococcus abyssi  
 C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
 C/Accession: F75103  
 R/Anonymous, Genoscope  
 submitted to the EMBL Data Library, July 1999  
 A/Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure  
 A/Reference number: A75001  
 A/Accession: F75103  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-880 <KAM>  
 A/Cross-references: UNIPROT:O5U2C8; GB:A1248286; GB:AL096836; NID:G5456366; PIDD:CA5013  
 A/Experimental source: strain Orsay  
 C/Genetics:  
 A/Genes: PAB0812  
 C/Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match  
 Best Local Similarity 11.0%; Score 115.5; DB 2; Length 880;  
 Matches 58; Conservative 38; Mismatches 78; Indels 67; Gaps 10;

QY 3 KKKGL-----SAEERTMMEIFSETQVDFOLDEKIAPKRGITANSVKEVLOS 53  
 DB 257 RKGLBEKIVQIERSEKKAISELEIVDIPKQEKERKLGFR----- 306  
 QY 54 LVDDGMVDCERIGTSNYWAFPSKAL-----HARKLEVLESQSGSKHSL--- 103  
 DB 307 --DEYSEKLRLEKELSMESSELKALIEVKEEKKERAEIRKLESEIRLEELKPY 364  
 QY 104 -----OKSIEKAK-----IGRCET--EERTLAKESSL-----RD 132  
 DB 365 VEEEDAKOVQKQERLAKLKSISPEVIEKLESLEKERTIEALKEITTRIGQNEGE 424  
 QY 133 QREOLKA--EVEKKKCDPQVEEIRQANKVAKKANRNTNIPATSMARKFGEENK 190  
 DB 425 KNEBKAIIEIRKAKGKCPVCGELTEBHK--KELMEVYILEIKIEIEELKRTTE--EERK 481  
 QY 191 I 191  
 DB 482 L 482

## RESULT 10

S24972  
 tropomyosin alpha, cardiac - pig  
 C/Species: Sus scrofa domestica (domestic pig)  
 C/Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
 C/Accession: S24972  
 R/Whitby, F.G.; Kent, H.M.; Stewart, F.; Stewart, M.; Xie, X.; Hatch, V.; Cohen, C.; Phil  
 submitted to the EMBL Data Library, April 1992  
 A/Description: Structure of tropomyosin at 9 Angstroms resolution.  
 A/Reference number: S24972  
 A/Accession: S24972  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-284 <WHI>  
 A/Cross-references: UNIPROT:P42639; EMBL:X66274; NID:G1926; PIDD:CAA46966.1; PID:G1927  
 C/Superfamily: tropomyosin  
 C/Keywords: cardiac muscle; heart

Query Match  
 Best Local Similarity 23.2%; Pred. No. 1.3; Length 284;  
 Matches 52; Conservative 37; Mismatches 61; Indels 74; Gaps 11;

QY 1 MSKKGLSAE---EKTRMMEIFSETQVDFOLDEKIAPKRGITANSVKEVLOS 57  
 DB 73 LAEKKADTAEADVASINRRIOLEFEELDRAQ---ERLA-----TALQLEAEKRADE 122  
 QY 58 --GMVDCERIGTSNYWAFPSKALHAR---KHLLEVLESQSGSKHSL----- 100

DB

123 SERGM-----

-----KVESRAQDEKKEHIOELQKEA--KHIAEDARKYE 163

QY

101 -----

-----ASLQKSIIEKAKI--GRCTEERTRLAKELSSLRDREOLKAEVRYDCD 148

DB

164 EVARLVIIESDLERAERAELESGKC-----AELEBELKVTNNKLSLEAKYQKE 218

QY

149 PQVEEIRQANKVAKKANRNTNIPATSMARKFGEENKID 192

DB

219 DVEEELKIVLSDIKLEAEFR-----AEFRKSVTKLEKSID 254

## RESULT 11

T05409  
 hypothetical protein F106.170 - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
 R/Bevan, M.; Weichselgartner, M.; Fartmann, B.; Granderath, K.; Dauner, D.; Herzl, A  
 submitted to the Protein Sequence Database, February 1998  
 A/Reference number: Z15414  
 A/Accession: T05409  
 A/Molecule type: DNA  
 A/Residues: 1-764 <BEV>  
 A/Cross-references: UNIPROT:O49371; EMBL:AL021811  
 A/Experimental source: cultivar Columbia; BAC clone F106  
 A/Map position: 4  
 A/Note: F106.170

Query Match  
 Best Local Similarity 10.9%; Score 114; DB 2; Length 764;  
 Matches 42; Conservative 40; Mismatches 78; Indels 24; Gaps 5;

QY 8 SAEERTMMEIFSETQVDFOLDEKIAPKR-----GITANSVKEVLOS 57  
 DB 49 SAESANVLPDKLFAKTHLEKOTQHSVYDDDLVSNLGVLESDELPAALVALKEED 108  
 QY 59 MVDCEKIGTSNYWAFPSKALHARKKLEVLESQSGSKHSLQKSIIEKAKI--ELA 159  
 DB 109 LMDERKILSD-----KNKLNKAKELERKRTTISEASLHSEISQELKRAV--- 178  
 QY 119 ERTLAKESSLRDREOLKAEVRYDCDQVVEEIRQANKVAKKANRNTNIPATSM 178  
 DB 160 SQAREIEELGHKLRERDERALQSSLTLEKELEKRYQ-----ELANSKEVSMALIS 213  
 QY 179 WAKR 182  
 DB 214 FEKK 217

## RESULT 12

A32183  
 tropomyosin TPM1 - yeast (Saccharomyces cerevisiae)  
 N/Alternate names: Protein N2132; protein YNL079c  
 C/Species: Saccharomyces cerevisiae  
 C/Date: 08-Sep-1989 #sequence\_revision 08-Sep-1989 #text\_change 09-Jul-2004  
 R/Jin, H.; Breitscher, A.  
 Cell 57, 233-242, 1989  
 A/Title: Disruption of the single tropomyosin gene in yeast results in the disappearance  
 A/Reference number: A32183; MVID:89195234; PMID:2649250  
 A/Accession: A32183  
 A/Molecule type: DNA  
 A/Residues: 1-199 <LTU>  
 A/Cross-references: UNIPROT:P17536; EMBL:M25501; NID:G173037; PIDD:AAA35174.1; PID:G1730  
 submitted to the EMBL Data Library, April 1995  
 A/Reference number: S53896  
 A/Accession: S53896  
 A/Molecule type: DNA  
 A/Residues: 1-199 <POB>  
 A/Cross-references: EMBL:X66470; NID:G791101; PIDD:CAA60179.1; PID:G791105  
 R/Boehlmann, R.; Philippen, P.

submitted to the Protein Sequence Database, April 1996  
 A:Reference number: S62997  
 A:Accession: S63011  
 A:Molecule type: DNA  
 A:Residues: 1-199 <POW>  
 A:Cross-references: EMBL:Z71355; NID:g1301970; PIDN:CAA95953.1; PID:g1301971; MIPS:YNL07  
 A:Experimental source: strain S288C  
 R:Solier-Mira, A.; Saitz, J.E.; Ballesta, J.P.G.; Remacha, M.  
 submitted to the Protein Sequence Database, April 1996  
 A:Reference number: S63018  
 A:Accession: S63018  
 A:Molecule type: DNA  
 A:Residues: 1-199 <SOI>  
 A:Cross-references: EMBL:Z71355; NID:g1301970; PIDN:CAA95953.1; PID:g1301971; MIPS:YNL07  
 A:Experimental source: strain S288C  
 R:Boehlmann, R.; Philippesen, P.  
 Yeast 12, 391-402, 1996  
 A:Title: Sequencing a cosmid clone of Saccharomyces cerevisiae chromosome XIV reveals 12  
 A:Reference number: S63925; MUID:96267764; PMID:8701611  
 A:Accession: S63928  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-199 <POF>  
 A:Cross-references: EMBL:X86470; NID:g791101; PIDN:CAA60179.1; PID:g791105  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995  
 C:Genetics:  
 A:Gene: SGD.TPM1  
 A:Cross-references: SGD:S0005023; MIPS:YNL079C  
 A:Map position: 14L  
 C:Superfamily: tropomyosin TPM1  
 C:Keywords: coiled coil; cytoskeleton

Query Match 10.8%; Score 113; DB 2; Length 199;  
 Best Local Similarity 24.7%; Pred. No. 1;  
 Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;  
 QY 7 LSAERKTRMMEIFSETKQVFPQKLEKIAPEKGTATSVESVLQSLVDD-----GMV 60  
 DB 13 LEAEWQGEKVELEKKNQDLQ-ENVE-----KENQKSLTVN-QLQLEDELEKLEAGLS 65  
 QY 61 DCRIGTSNYWAFPSKALHARKHL-----EVLQSLSG-----SQK 99  
 DB 66 DSKQREODVKEKQIKSLTVNQHLEBEIEKLEAEALSKQLSEDSHHLQSNNDNFSKX 125  
 QY 100 HASLQKSLFKAKIGCETFEFRR-----LAKELSLRDQREOLKAVE-----YKRC 147  
 DB 126 NQOLEDELESDPTKLETKLEKLESDLKADQLERRVALEQREWERKNEELTVYCYEDA 185  
 QY 148 DPQVVE 153  
 DB 186 KKELE 191

RESULT 13  
 JG6199  
 alpha-tropomyosin S-1 - axolotl  
 C:Species: Ambystoma mexicanum (axolotl)  
 C:Date: 11-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 09-Jul-2004  
 C:Accession: JG6199  
 R:Luque, E.A.; Spinner, B.J.; Dube, S.; Dube, D.K.; Lemanski, L.F.  
 Gene 185, 175-180, 1997  
 A:Title: Differential expression of a novel isoform of alpha-tropomyosin in cardiac and  
 A:Reference number: JG6198; MUID:97208870; PMID:9055812  
 A:Contents: skeletal muscle  
 A:Accession: JG6199  
 A:Molecule type: mRNA  
 A:Residues: 1-284 <LUO>  
 A:Cross-references: UNIPROT:P87349; GB:U13450; NID:g1871357; PIDN:AA60092.1; PID:g18713  
 C:Comment: This protein is a actin-binding protein.  
 C:Genetics:  
 A:Gene: Atms-1  
 C:Superfamily: tropomyosin  
 C:Keywords: actin binding

Query Match 10.7%; Score 112; DB 2; Length 284;  
 Best Local Similarity 22.7%; Pred. No. 1.8;  
 Matches 48; Conservative 41; Mismatches 74; Indels 48; Gaps 9;  
 QY 1 MSKKGLSAEE---KRTMMEIFSETKQVFO-----LKQLEKIAPEK-EGKGTANSV 47  
 DB 73 LARKKATPAESDVASLNRIRQLVEEELRPAERLATALQKEAEKADDESERGMKVIEN 132  
 QY 48 KEVLSLVDDGVDCERIGTSNYWAFPSKALHARKHLEVLQSLQSGSQK-----HASL 103  
 DB 133 R---ALQDEKEMLEQEI-----QLQEKHIAEADRYEEVARKLVITIEGDL 176  
 QY 104 QKSLKAKI---GRCTEERTRLAKELSLRDQREOLKAVEKYKDCDPQVVEIRQANKV 161  
 DB 177 ERAERARLSEKQK-----AELEELKTVTNLKLAEQAKEYSQKEDKYEELIVLTDK 231  
 QY 162 AKKANRWTDNIPAIKSWAKRKFGEENKID 192  
 DB 232 LKEAETR-----AEFAERTVAKLEKSID 254

RESULT 14  
 A24199  
 tropomyosin NM, skeletal muscle - human  
 C:Species: Homo sapiens (man)  
 C:Date: 02-Jun-1998 #sequence\_revision 02-Jun-1998 #text\_change 13-Aug-1999  
 C:Accession: S06210; A24199  
 R:Clayton, L.; Reinach, F.C.; Chumbley, G.M.; Macleod, A.R.  
 J. Mol. Biol. 201, 507-515, 1988  
 A:Title: Organization of the hTm(m) gene. Implications for the evolution of muscle and  
 A:Reference number: S02554; MUID:88332987; PMID:3418707  
 A:Accession: S06210  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-285 <CLA>  
 R:Reinach, F.C.; Macleod, A.R.  
 Nature 332, 648-650, 1986  
 A:Title: Tissue-specific expression of the human tropomyosin gene involved in the genere  
 A:Reference number: A24199; MUID:86311274; PMID:3015581  
 A:Accession: A24199  
 A:Molecule type: mRNA  
 A:Residues: 1-285 <REI>  
 A:Cross-references: GB:X04201; NID:g37429; PIDN:CAA27798.1; PID:g37430  
 A:Note: an intronless pseudogene resembling this mRNA is also known  
 C:Superfamily: tropomyosin  
 C:Keywords: alternative splicing; coiled coil; muscle; skeletal muscle

Query Match 10.7%; Score 112; DB 2; Length 285;  
 Best Local Similarity 23.5%; Pred. No. 1.8;  
 Matches 52; Conservative 41; Mismatches 72; Indels 56; Gaps 9;  
 QY 1 MSKKGLSAEE---KRTMMEIFSETKQVFO-----LKQLEKIAPEK-EGKGTANSV 47  
 DB 74 LAERKADAEAEVAVSLNRIRQLVEEELRPAERLATALQKEAEKADDESERGMKVIEN 133  
 QY 48 KEVLSLVDDGVDCERIGTSNYWAFPSKALHARKHLEVLQSLQSGSQK-----HASL 107  
 DB 134 R---ALQDEKEMLEQEI-----QLQEKHIAEADRYEEVARKLVITIEGDL 177  
 QY 108 EKAKIGRCETERTTLA-----KELSLRDQREOLKAVEKYKDCDPQVVEIRQANKV 159  
 DB 178 ER-----TERRALASKSCLELEELKVTNNLKLAEQAKEYSQKEDKYEELIKILT 230  
 QY 160 KVAKKANRWTDNIPAIKSWAKRKFGEENKIDRTFGIPED 200  
 DB 231 DLKKEAETRAE---PAERSVA-----KLEKTIIDLED 259

RESULT 15  
 S49143  
 Eg10 protein - tapeworm (Echinococcus granulosus)  
 C:Species: Echinococcus granulosus



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: March 23, 2005, 19:08:01 ; Search time 177 Seconds

(without alignments)  
593.086 Million cell updates/sec

Title: US-10-087-190-3

Perfect score: 1047  
Sequence: 1 MSKKKGSLAEKRTIMEIF.....FEENKIDRTFGIPEDFDYID 205

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1047	100.0	205	Q9BWT6	Q9BWT6 homo sapien
2	954	91.1	205	Q8K396	Q8K396 mus musculu
3	948	90.5	205	Q9D0A1	Q9D0A1 mus musculu
4	644	61.5	220	Q6DC61	Q6DC61 brachydanio
5	462	44.1	196	Q8EEZ8	Q8EEZ8 schistosoma
6	440	42.0	230	Q8GYD2	Q8GYD2 arabidopsis
7	427	40.8	207	Q6H432	Q6H432 oryza sativ
8	349	33.3	210	YAS3_SCHPO	Q09739 schizosacch
9	340.5	32.5	203	Q8SU9	Q8SU9 encephalito
10	340.5	32.5	203	Q6WDA3	Q6WDA3 giardia lam
11	340.5	32.5	203	Q7QTX0	Q7QTX0 giardia lam
12	332.5	31.8	179	Q6Q9F9	Q6Q9F9 aedes aegy
13	277	26.5	196	Q7RH53	Q7RH53 plasmodium
14	266.5	25.5	211	Q6BOL5	Q6BOL5 debaromyce
15	221	21.1	225	Q7SCM5	Q7SCM5 ashyia goss
16	206.5	19.7	225	Q6CSX5	Q6CSX5 kluyveromyc
17	197	18.8	128	Q9S2B5	Q9S2B5 arabidopsis
18	175	16.7	222	Q6FL56	Q6FL56 candida gla
19	168.5	16.1	174	YGR3_YEAST	YGR3_YEAST saccharomy
20	153.5	14.7	101	Q6E679	Q6E679 nosema locu
21	136.5	13.0	910	Q7T2F8	Q7T2F8 brachydanio
22	135.5	12.9	1281	Q6AWE1	Q6AWE1 mus musculu
23	134.5	12.8	890	Q6AWB1	Q6AWB1 homo sapien
24	134.5	12.8	890	Q6AWB3	Q6AWB3 homo sapien
25	134.5	12.8	1139	Q6IQ37	Q6IQ37 homo sapien
26	134.5	12.8	1264	Q6NZM3	Q6NZM3 mus musculu
27	134.5	12.8	1278	Q6MZ23	Q6MZ23 homo sapien
28	134.5	12.8	1278	Q6MZ23	Q6MZ23 homo sapien
29	131.5	12.6	1280	Q6MZ23	Q6MZ23 rattus norv
30	130.5	12.5	1224	Q6PCJ1	Q6PCJ1 xenopus lae
31	126.5	12.1	1232	Q6PCJ1	Q6PCJ1 xenopus lae

32	126	12.0	609	2	Q8TXA4	Q8TXA4 mechanopyru
33	122.5	11.7	797	2	Q6GLB3	Q6GLB3 xenopus tro
34	122	11.7	10578	2	Q8ISF5	Q8ISF5 caenorhabdi
35	122	11.7	18519	2	Q8ISF6	Q8ISF6 caenorhabdi
36	122	11.7	18534	2	Q8ISF7	Q8ISF7 caenorhabdi
37	121.5	11.6	995	2	Q6FIN4	Q6FIN4 meoplasma
38	121.5	11.6	1357	1	KTNI_HUMAN	KTNI_HUMAN
39	121.5	11.6	1364	1	KTNI_CHICK	KTNI_CHICK
40	120.5	11.5	172	2	Q824V3	Q824V3 gallus gall
41	120.5	11.5	448	2	Q84H59	Q84H59 anaplasm m
42	120.5	11.5	1732	2	Q9YJ35	Q9YJ35 drosophila
43	120	11.5	882	1	RA50_PYRPU	RA50_PYRPU
44	120	11.5	1120	1	ERCL_MOUSE	ERCL_MOUSE
45	118.5	11.3	1177	2	Q87711	Q87711 pyrococcus

## ALIGNMENTS

RESULT 1  
Q9BWT6 PRELIMINARY; PRT: 205 AA.  
ID Q9BWT6  
AC Q9BWT6; 01-JUN-2001 (TREMblrel. 17, Created)  
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
DT 25-OCT-2004 (TREMblrel. 28, Last annotation update)  
DE GAF.  
GN Name=GAF;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Solis G., Hofer H.W.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymph;  
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strusberg R.U., Reingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein W.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.,  
RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Bosak S.A., Loquellano N.A., Peters G.J., Abrahamson R.D., Mullahy S.J.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,  
RA Krzywinski M.I., Skalski U., Smallos D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Maira M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences".  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymph;  
RA Strusberg R.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY028916; AAK26168.1; -;  
DR EMBL; BC032142; AAH32142.1; -;  
DR InterPro; IPR005647; Mnd1.  
DR InterPro; IPR009058; wing\_hlx\_DNA\_bnd.  
DR Pfam; PF03962; Mnd1; 1.  
SQ SEQUENCE 205 AA; 23753 MW; 95B0B14068DA0B51 CRC64;

Query Match 100.0%; Score 1047; DB 2; Length 205;

Best Local Similarity 100.0%; Pred. No. 5.9e-56;  
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSKKKGLSAEKRTRMEIFSETKDVFOQLDEKLAPEKGTITANSVEVLOSLVDDGMV
DB 1 MSKKKGLSAEKRTRMEIFSETKDVFOQLDEKLAPEKGTITANSVEVLOSLVDDGMV
QY 61 DCEIRGTSNYWMAFPSPALHARKHLEVEESQSSGSHASLOKSEIKAKIGRCETEER
DB 61 DCEIRGTSNYWMAFPSPALHARKHLEVEESQSSGSHASLOKSEIKAKIGRCETEER
QY 121 TRIAKELSLRDRQOLKAEVEKRCDDPOVVEEIRQANKYAKKAAANRWTDNIFAIKSWA
DB 121 TRIAKELSLRDRQOLKAEVEKRCDDPOVVEEIRQANKYAKKAAANRWTDNIFAIKSWA
QY 121 TRIAKELSLRDRQOLKAEVEKRCDDPOVVEEIRQANKYAKKAAANRWTDNIFAIKSWA
DB 121 TRIAKELSLRDRQOLKAEVEKRCDDPOVVEEIRQANKYAKKAAANRWTDNIFAIKSWA
QY 181 KRKFGFENKIDRTFGIPEDFDYID 205
DB 181 KRKFGFENKIDRTFGIPEDFDYID 205

```

## RESULT 2

```

ID 08K396 PRELIMINARY; PRT; 205 AA.
AC 08K396;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE GAV protein.
GN Name=2610034E18R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary tumor;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603999;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Altschul S.F., Zeeberg B., Buelow K.H., Schenfer C.M., Schuler G.D.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Franke C.,
RA Raha S.S., Lounellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Abramson R.D., Mullaly S.J.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
RA Kizylinski M.I., Skalka U., Smalish D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary tumor;
RX Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027741; AAH27741.1;
DR MGI; MGI:1924165; 2610034E18R1k.
DR InterPro; IPR005647; Mnd1.
DR Pfam; PF03962; Mnd1; 1.
SQ SEQUENCE 205 AA; 23849 MW; 122C3FA9B4325120 CRC64;

```

Query Match 91.1%; Score 954; DB 2; Length 205;

Best Local Similarity 89.8%; Pred. No. 2.5e-50;

Matches 184; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

1 MSKKKGLSAEKRTRMEIFSETKDVFOQLDEKLAPEKGTITANSVEVLOSLVDDGMV 60

```

DB 1 MSKKKGLSAEKRTRMEIFSETKDVFOQLDEKLAPEKGTITANSVEVLOSLVDDGMV
QY 61 DCEIRGTSNYWMAFPSPALHARKHLEVEESQSSGSHASLOKSEIKAKIGRCETEER
DB 61 DCEIRGTSNYWMAFPSPALHARKHLEVEESQSSGSHASLOKSEIKAKIGRCETEER
QY 121 TRIAKELSLRDRQOLKAEVEKRCDDPOVVEEIRQANKYAKKAAANRWTDNIFAIKSWA
DB 121 TRIAKELSLRDRQOLKAEVEKRCDDPOVVEEIRQANKYAKKAAANRWTDNIFAIKSWA
QY 121 TRIAKELSLRDRQOLKAEVEKRCDDPOVVEEIRQANKYAKKAAANRWTDNIFAIKSWA
DB 121 TRIAKELSLRDRQOLKAEVEKRCDDPOVVEEIRQANKYAKKAAANRWTDNIFAIKSWA
QY 181 KRKFGFENKIDRTFGIPEDFDYID 205
DB 181 KRKFGFENKIDRTFGIPEDFDYID 205

```

## RESULT 3

```

ID 09DOA1 PRELIMINARY; PRT; 205 AA.
AC 09DOA1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
DE enriched library, clone:2610034E18 product:GAV homolog.
GN Name=2610034E18R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085600; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Komori H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1517-1530(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kaishiwagi K.,
RA Yoneda S., Inoue K., Togawa K., Izawa M., Ohara E., Watanabe K.,
RA Taniwaki Y., Ishikawa T., Ozawa K., Tanaka T., Matsuzura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format

```



[illegible]

```

Db      :|||||
        62 FMFPPSKAOKRNRIIEKTDGIDHTNRQIFKTRTSINELSKRKOTIEENKRLINLET
Qy      131 RDQREQLKAEVEKYKDCDPQVEEIRQANKVAEANKRTWDNI.FALKSNAPKRGFEENK 121
Db      122 KILBELTALHEDODKEDHPDRLSEIRQOOLVADSNARKMTDNI.FIKSLNKSFLDEAT 190
Qy      191 IIRTGIPEDPDYI 204
Db      182 FCRPEIPENFDYI 195

```

## RESULT 6

```

O6GYD2  PRELIMINARY; PRT; 230 AA.
ID  O6GYD2
AC  O6GYD2;
DT  01-MAR-2003 (TREMBlrel. 23, Created)
DT  05-JUL-2004 (TREMBlrel. 23, last sequence update)
DE  Hypothetical protein At4g29170/F19B15_200 (Hypothetical protein
DE  At4g29170).
GN  Name=At4g29170/F19B15_200; Synonyms=At4g29170;
OS  Arabidopsis thaliana (mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; euclcotyledons; core eudicots; rosids;
OX  NCBI_TaxId=3702;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA  Nakajima M., Enju A., Kamiya A., Natsuka M., Carninci P., Kawai J.,
RA  Hayashizaki Y., Shinozaki K.,
RL  Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Yamada K., Chan M.M., Chang C.H., Dale J.M., Heuan V.W., Lee J.M.,
RA  Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
RA  Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA  Ishida J., Jones T., Kamiya A., Kawai J., Kim C.U., Natsuka M.,
RA  Nishiyama M., Palm C.J., Sakurai T., Satou M., Seki M., Shimizu P.,
RA  Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA  Theologis A.,
RL  Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AK117113; BAC42364.1;
DR  EMBL; BT005435; AAO63855.1;
DR  InterPro; IPR005639; endotoxin_N.
DR  InterPro; IPR005647; Mnd1.
DR  Pfam; PF03962; Mnd1.1.
KW  Hypothetical protein.
SQ  SEQUENCE 230 AA; 26402 MW; 131BBS146D8C91ED CRC64;

```

Query Match 42.0%; Score 440; DB 2; Length 230;

Best Local Similarity 43.8%; Pred. No. 3.2e-19;

Matches 91; Conservative 43; Mismatches 70; Indels 4; Gaps 2;

```

Qy      1 MSKKKGSAEKKRTMMEIFSETKVDFOLKDEKIAPEKKITMSYKVLQSLVDGMY 60
        |||||
Db      1 MSKKKGSLSEKREKMLQIFYSQDFILKELKKGPR-KGVISQSVKVDVIGSLVDDIV 59
Qy      61 DCRIGTSNYVAFPSKALHARKHLEVLSELSQSGKASLOKSIEXKIGRCCTER 120
        :|||
Db      60 ANDKIGISIFWLSLPSGANGNRSVRQKLESDDGSKNRKLAELVDCEMLKKGRESEER 119
Qy      121 TLAKELSLRQREQLKAEVEKYKDCDPQVEEIRQANKVAEANKRTWDNI.FALKSMA 180
        |
Db      120 TEALQOLDIEKKNHDLKNEWQFADNDPATLEKKNALFVHQSANRWTDNI.FILRQMC 179
Qy      181 KRKGFEENKIDRTF---GIPEPDYID 205
        :|||
Db      180 SNNFPQAEQLEHLYTAGITEDPDYIE 207

```

## RESULT 7

```

O6H432  PRELIMINARY; PRT; 207 AA.
ID  O6H432
AC  O6H432;
DT  05-JUL-2004 (TREMBlrel. 27, Created)
DT  05-JUL-2004 (TREMBlrel. 27, last sequence update)
DE  Putative GAI protein.
GN  Name=P0651G05.20;
OS  Oryza sativa (japonica cultivar-group).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Euphorbiales; Magnoliophyta; Liliopsida; Poales; Poaceae;
OX  NCBI_TaxId=39947;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Sasaki T., Matsumoto T., Katayose Y.;
RT  "Oryza sativa nippondate (GA3) genomic DNA, chromosome 9, PAC
RL  clone:P0651G05.2";
DR  EMBL; AP006528; BAD26517.1;
DR  InterPro; IPR005647; Mnd1.
DR  Pfam; PF03962; Mnd1.1.
SQ  SEQUENCE 207 AA; 23900 MW; 6A9C5EEB6D586E CRC64;

```

Query Match 40.8%; Score 427; DB 2; Length 207;

Best Local Similarity 42.5%; Pred. No. 1.7e-18;

Matches 88; Conservative 44; Mismatches 71; Indels 4; Gaps 2;

```

Qy      1 MSKKKGSAEKKRTMMEIFSETKVDFOLKDEKIAPEKKITMSYKVLQSLVDGMY 60
        |||||
Db      1 MSKKKGSLSEKREKMLQIFYSQDFILKELKKGPR-KGVISQSVKVDVIGSLVDDIV 59
Qy      61 DCRIGTSNYVAFPSKALHARKHLEVLSELSQSGKASLOKSIEXKIGRCCTER 120
        |||||
Db      60 IKKIKTSYVFWLSLPSGANGNRSVRQKLESDDGSKNRKLAELVDCEMLKKGRESEER 119
Qy      121 TLAKELSLRQREQLKAEVEKYKDCDPQVEEIRQANKVAEANKRTWDNI.FALKSMA 180
        :|||
Db      120 EALBELRAVQNHKKKEELAAVADSDPALAMMDALFVHQAANRWTDNI.FILQMC 179
Qy      181 KRKGFEENKIDRTF---GIPEPDYI 204
        :|||
Db      180 STTFQAEQLEHLYTAGITEDPDYIE 206

```

## RESULT 8

YAS3 SCHPO

ID YAS3 SCHPO STANDARD; PRT; 210 AA.

```

AC  009719;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, last sequence update)
DE  Hypothetical protein C13A11.03 in chromosome 1.
GN  ORFNames=SPAC13A11.03;
OS  Schizosaccharomyces pombe (fission yeast).
OC  Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC  Schizosaccharomycetales; Schizosaccharomycetaceae;
OX  NCBI_TaxId=4896;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  STRAIN=972;
RX  MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA  Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA  Sgourou J., Peat N., Hayes D., Baker S., Basham D., Bowman S.,
RA  Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA  Collins M., Connor R., Cronin A., Davis P., Fellows I., Fraser A.,
RA  Gentile S., Goble A., Hamlin N., Harris D., Hildalgo J., Hodgson G.,
RA  Holroyd S., Hornsby T., Howarth S., Huckle E.D., Hunt S., Jagels K.,
RA  James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA  Mooney P., Moulé S., Mungall K., Murphy L., Niblett D., Odell C.,
RA  Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA  Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

```

RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Meller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambach R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,  
 RA Ceretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Usero D., Barrett B.G., Nurse P.,  
 RT "The genome sequence of *Schizosaccharomyces pombe*,"  
 RL Nature 415:871-880 (2002).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL, 254096; CAA90804.1; -  
 DR PIR, T37610; T37610.  
 DR GeneDB, Spombe: SPAC13A11.03; -  
 DR InterPro, IPR005647; Mnd1.  
 DR InterPro, IPR009058; Wing\_hlx\_DNA\_bnd.  
 DR Pfam, PF03962; Mnd1; 1.  
 DR Hypothetical protein.  
 SQ SEQUENCE 210 AA; 24224 MW; FA546F070A37665 CRC64;  
 Query Match 33.3%; Score 349; DB 1; Length 210;  
 Best Local Similarity 41.6%; Pred. No. 9e-14;  
 Matches 84; Conservative 36; Mismatches 76; Indels 6; Gaps 5;  
 QY 5 KGLSAEKRTRMEIFSETKDVFLQKDLKIAPEKGIITAMSVKVELQSLVDGMDGDCER 64  
 DB 4 KGLSLAEKRRRLAEALFHDSKDFQLKEVEKLSK-KQIVLTQTKVDLOSVDNDNIVKTEK 62  
 QY 65 IGTSNYVAFPKALHARKHKLVLSEQLSESGQKASLQKSI--EKAKIGRCETE-EKT 121  
 DB 63 IGTSNYVSPFSDAKRSNRSVSLQAOQLDLKQKSKTLDENISPEKSRKDNHGTENDAN 122  
 QY 122 RLAKELSLRD-QREQLKAEVEKYKDCDPQVEEIRQANKVAKAANRWTNIPAIKSWA 180  
 DB 123 QYTLLELHAKESSELKLTQLSLNLHNCNPFELKKNENTKYWEAANLMTDQIHLLIAPC 182  
 QY 181 KRKPFENKIDRTGIPEDFD 202  
 DB 183 -RDMGADTNQIREYCSIPEDLD 203  
 RESULT 9  
 Q8SU9 PRELIMINARY; PRT; 203 AA.  
 AC Q8SU9;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Hypothetical protein ECU10\_1600.  
 OS Name=ECU10\_1600;  
 OS Eucaryota; Fungi; Microsporidia; Unikaryoniidae; Eucephalitozoon.  
 OC NCB1\_Taxid=284813;  
 OX NCB1\_Taxid=284813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB-M1;  
 RX MEDLINE=21576510; PubMed=11719806; DOI=10.1038/35106579;  
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,  
 RA Prensier G., Barde V., Peyretailade E., Broctier P., Wincker P.,

RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,  
 RA Weissenbach J., Vivares C.P.,  
 RT "Genome sequence and gene compaction of the eukaryote parasite  
 RT *Eucephalitozoon cuniculi*,"  
 RL Nature 414:450-453 (2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB-M1;  
 RA Genoscope;  
 RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL, AL590448; CAD25881.1; -  
 DR InterPro, IPR002114; HPr\_Serp\_S.  
 DR InterPro, IPR005647; Mnd1.  
 DR InterPro, IPR009058; Wing\_hlx\_DNA\_bnd.  
 DR Pfam, PF03962; Mnd1; 1.  
 DR PROSITE, PS00589; PTS\_HPR\_SER; UNKNOWN\_1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 203 AA; 23743 MW; DFE33A65A1A28A42 CRC64;  
 Query Match 32.5%; Score 340.5; DB 2; Length 203;  
 Best Local Similarity 35.1%; Pred. No. 2.8e-13;  
 Matches 71; Conservative 50; Mismatches 72; Indels 9; Gaps 3;  
 QY 7 LSAEKRTRMEIFSETKDVFLQKDLKIAPEKGIITAMSVKVELQSLVDGMDGDCERIG 66  
 DB 6 MSLDQKSLILEIRIGSKSFYQLQELSLGSK-KGIIVNTIKELIQLVDLGLVTAKEVG 64  
 QY 67 TSNLYWAFPSKALHARKHKLVLSEQLSESGQKASLQKSI-EKAKIGRCETEERTLAE 126  
 DB 65 TSNLYWSPASBGIOKKRKRCLEMECERSODICKREKYEILENKSKRYTERELENK 124  
 QY 127 LSSL-----RDREQLKAEVEKYKDCDPQVEEIRQANKVAKAANRWTNIPAIKSWA 182  
 DB 125 LNALMKIEDQRE-----ELGKEERTPIAYDLVADRKEMDECRRIIDNVYIIQDYCS 180  
 QY 183 KRKPFENKIDRTGIPEDFDYI 204  
 DB 181 KRPFMEKSEFNSFGIPDLDI 202  
 RESULT 10  
 Q6MDA3 PRELIMINARY; PRT; 203 AA.  
 AC Q6MDA3;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Mnd1.  
 OS Giardia lamblia (Giardia intestinalis).  
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
 OX NCB1\_Taxid=5741;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ramesh M.A., Malik S.B., Logsdon J.M. Jr.,  
 RL Submitted (Apr-2003) to the EMBL/Genbank/DBJ databases.  
 DR EMBL, AY295092; AA024512.1; -  
 DR InterPro, IPR005647; Mnd1.  
 DR InterPro, IPR009058; Wing\_hlx\_DNA\_bnd.  
 DR Pfam, PF03962; Mnd1; 1.  
 SQ SEQUENCE 203 AA; 23278 MW; CC0625DCDC4158C0 CRC64;  
 Query Match 32.5%; Score 340.5; DB 2; Length 203;  
 Best Local Similarity 34.8%; Pred. No. 2.8e-13;  
 Matches 70; Conservative 46; Mismatches 84; Indels 1; Gaps 1;  
 QY 5 KGLSAEKRTRMEIFSETKDVFLQKDLKIAPEKGIITAMSVKVELQSLVDGMDGDCER 64  
 DB 4 KGTSLDEKERRLEBMLKRGKISNKTIELTS-KPTGISSWIKVVLQALVNEEDLVDTK 62  
 QY 65 IGTSNYVAFPKALHARKHKLVLSEQLSESGQKASLQKSI-EKAKIGRCETEERTLAE 124  
 DB 63 IGATSYWCFASKRSQAARTELARLOKALEGTNFIKATATARIELKVGRETERSSIL 122

QY 125 KELSRLDQREQLKAEVEKYKDCDPQVEEIRQANKVAKAEANRWTDNIFAIKSMARKKF 184  
 DB 123 KEXLALQVLEBQGRGTFRLDKNDPVAQKLRNTYDIKAGKANLMTDNIPLQKXYMTKL 182  
 QY 185 GFEEKKIDRTFGIEDPDYID 205  
 DB 183 QMDKKTVSTALGITGEFDYLE 203

## RESULT 11

Q7OTX0 PRELIMINARY; PRT; 203 AA.  
 AC 07OTX0:  
 DT 01-MAR-2004 (TREMBLrel. 26, Created)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE GLP\_76\_12374\_11763.  
 OS Giardia lamblia ATCC 50803.  
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
 OX NCBI\_TaxID=184922;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WB C6;  
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D., Olsen G.J., Sogin M.L.;  
 RT "Draft sequence of the Giardia lamblia genome.";  
 RI Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.  
 DR EMBL; AACB01000100; EAA38463.1; -.  
 DR InterPro; IPR005647; Mnd1.  
 DR Pfam; PF03962; Mnd1; 1.  
 DR Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 SQ SEQUENCE 203 AA; 23278 MW; CC0625DCD4158C0 CRC64;

Query Match 32.5%; Score 340.5; DB 2; Length 203;  
 Best Local Similarity 34.8%; Pred. No. 2.8e-13;  
 Matches 70; Conservative 46; Mismatches 84; Indels 1; Gaps 1;

QY 5 KGLSABEKRTRMEIFSETDVQPLDKLEKIAPEKGIATMSVKEVLQSLVDGMDGCR 64  
 DB 4 KGLSLEBKEKRLLEMLKGEIYSNKTIFLS-KPLGISSMVTKVQLVNVEDLDVDDK 62  
 QY 65 IGSNYYWAFPSKALARKHKLVELESQSEGSQKASLOKSIKAKIGRCETERTRLA 124  
 DB 63 IGSNYYWAFPSKALARKHKLVELESQSEGSQKASLOKSIKAKIGRCETERTRLA 122  
 QY 125 KELSRLDQREQLKAEVEKYKDCDPQVEEIRQANKVAKAEANRWTDNIFAIKSMARKKF 184  
 DB 123 KEXLALQVLEBQGRGTFRLDKNDPVAQKLRNTYDIKAGKANLMTDNIPLQKXYMTKL 182  
 QY 185 GFEEKKIDRTFGIEDPDYID 205  
 DB 183 QMDKKTVSTALGITGEFDYLE 203

## RESULT 12

Q609F9 PRELIMINARY; PRT; 179 AA.  
 AC 0609F9:  
 DT 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE MND1 domain containing protein.  
 OS Aedes aegypti (Yellowfever mosquito).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.  
 OX NCBI\_TaxID=7159;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Salivary gland;  
 RA Chandra P.K., Wikel S.K.;

RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY552061; AAS79347.1; -.  
 DR InterPro; IPR005647; Mnd1.  
 DR Pfam; PF03962; Mnd1; 1.  
 SQ SEQUENCE 179 AA; 20900 MW; 6E0016920350E476 CRC64;

Query Match 31.8%; Score 332.5; DB 2; Length 179;  
 Best Local Similarity 37.1%; Pred. No. 7.6e-13;  
 Matches 66; Conservative 52; Mismatches 57; Indels 3; Gaps 2;

QY 1 MSK-KKGLSABEKRTRMEIFSETDVQPLDKLEKIAPEKGIATMSVKEVLQSLVDGMDG 59  
 DB 1 MSKRKKGISAGEKKSIVLEIFHQSEKPYQLDVERIAKDKGLAEQVKEILQNLVBEGQ 60  
 QY 60 VDCENIGSNYYWAFPSKALARKHKLVELESQSEGSQKASLOKSIKAKIGRCETERT 119  
 DB 61 VETDKIGSQYYWFPAPKARKLKQVFEQLQKQEMEQSNDKIAELQKIGTKESQGSRR 120  
 QY 120 RTRLAKESLRDQREQLKAEVE--KYKDCDPQVEEIRQANKVAKAEANRWTDNIFA 175  
 DB 121 SSEWFKNTLKEKQKSLSKDKAKLKQSDQNSVDKXNRLPDLDHANRWGDTTYS 178

## RESULT 13

Q7RH53 PRELIMINARY; PRT; 196 AA.  
 AC 07RH53:  
 DT 01-MAR-2004 (TREMBLrel. 26, Created)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Homo sapiens GAD, putative.  
 GN Name=PY04140;  
 OS Plasmodium yoelii yoelii.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=73239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=17XNL;  
 RX PubMed=12368865; DOI=10.1038/nature01099;  
 RA Carlson J.M., Angiolini S.V., Suh B.B., Kool J.T.W., Petrea M., Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L., Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L., Shalton S.J., van Aken S.B., Riedmiller S.B., Feldblum T.V., Cho J.K., Quackenbush J., Sedegah M., Shoabli A., Cummings L.M., Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A., Cunningham D.A., Preiser P.R., Bergman L.W., Valdivia A.B., van Lin L.H., Janse C.J., Waters A.P., Smith H.O., Gardner M.J., Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J., Carucci D.J.;  
 RT "Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii.";  
 RL Nature 419:512-519(2002).  
 CC -1- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.  
 DR EMBL; AABL01001237; EAA15959.1; -.  
 DR InterPro; IPR005647; Mnd1.  
 DR Pfam; PF03962; Mnd1; 1.  
 SQ SEQUENCE 196 AA; 23066 MW; 7FD8820707329074 CRC64;

Query Match 26.5%; Score 277; DB 2; Length 196;  
 Best Local Similarity 29.7%; Pred. No. 1.9e-09;  
 Matches 60; Conservative 52; Mismatches 76; Indels 14; Gaps 3;

QY 3 KKKGLSABEKRTRMEIFSETDVQPLDKLEKIAPEKGIATMSVKEVLQSLVDGMDG 62  
 DB 2 KKKGKSNEDKLLVDIMLESSEFLKELALAPK-KGIRSIYVKDLIQOUIDNKKKS 60  
 QY 63 ERIGSNYYWAFPSKALARKHKLVELESQSE-----GSQKASLOKSIKAKIGRC 115  
 DB 61 EKVGQNVFWILKTESSILQNKQYELDKKREYEMQAQKENTAELENSIS-----L 114  
 QY 116 ETEERTLAKESLRDQREQLKAEVEKYKDCDPQVEEIRQANKVAKAEANRWTDNIFA 175

```

Db      115 KTDKDKDTLKEVKNVLDISIEIKSELDKDKTDHIOIEKIKQSNFATESIERMNNIPL 174
Qy      176 IKSMAKKKGFEENKIDRTFGI 197
Db      175 LKQWIDRTKNSGDVDRLLGM 196

RESULT 14
06B0L5
ID      06B0L5      PRELIMINARY;      PRT;      211 AA.
AC      06B0L5;
DT      25-OCT-2004 (TREMBlrel. 28, Created)
DT      25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT      25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE      Similar to CA34131|PF9239 Candida albicans IPF9239 unknown function
       (Fragment).
GN      ORFNames=DEHA0B04994g;
OS      Debaryomyces hansenii CBS767.
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX      NCBI_TaxID=284592;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CBS767;
RG      Genolevures;
RA      Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
       Lafontaine I., de Montigny J., Marek C., Neuvéglise C., Talla E.,
       Goffard N., Frangul L., Aigle M., Anhouard V., Babour A., Barbe V.,
       Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
       Boissier A., Boyer J., Catolico L., Confantolero F., de Darvar A.,
       Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Gropi A.,
       Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
       Kerrest A., Kozulski R., Lemaire M., Lesur I., Ma L., Muller H.,
       Nicard J.M., Nikolaki M., Oztas S., Ozier-Kalogeropoulos O.,
       Pellenn S., Peltier S., Richard G.F., Straub M.L., Suleau A.,
       Swenne S., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
       Zenou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
       Bouchier C., Caudon B., Scarpelli C., Gaillardin C., Weissenbach J.,
       Wincker P., Souciet J.L.;
       "Genome evolution in yeasts.";
       Nature 430:35-44(2004).
       [2]
RA      SEQUENCE FROM N.A.
RC      STRAIN=CBS767;
RG      Genoscope;
RA      Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL; CR382137; CAG87731.1; -.
DR      InterPro; IPR005647; Mnd1.
DR      Pfam; PF03962; Mnd1; 1.
FT      NON_TER
SQ      SEQUENCE      211 AA; 24831 MW; F8B5DC6618334C8 CRC64;

Query Match      25.5%; Score 266.5; DB 2; Length 211;
Best Local Similarity 29.0%; Pred. No. 8.8e-09;
Matches 61; Conservative 55; Mismatches 75; Indels 19; Gaps 5;

```

```

RESULT 15
07SCH5
ID      07SCH5      PRELIMINARY;      PRT;      225 AA.
AC      07SCH5;
DT      05-JUL-2004 (TREMBlrel. 27, Created)
DT      05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT      05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE      ACT056Cp.
GN      ORFNames=ACT056C;
OS      Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX      NCBI_TaxID=33169;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 10895;
RA      Brachat S., Voegelé S.E., Dietrich F.S., Lerch A., Gaffney T.,
       Philippsen P.;
       Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AB016816; AAS51172.1; -.
DR      AGD; ACT056C; -.
DR      InterPro; IPR005647; Mnd1.
DR      Pfam; PF03962; Mnd1; 1.
SQ      SEQUENCE      225 AA; 25753 MW; BA73C889FFBEB90F CRC64;

Query Match      21.1%; Score 221; DB 2; Length 225;
Best Local Similarity 26.4%; Pred. No. 5.3e-06;
Matches 56; Conservative 51; Mismatches 93; Indels 12; Gaps 5;

```

Search completed: March 23, 2005, 19:17:52  
Job time : 180 secs

**THIS PAGE BLANK (USPTO)**